

Instituto Tecnológico de Costa Rica

Escuela de Ingeniería en Electrónica



Reconocimiento de patrones acústicos para la protección del ambiente utilizando *wavelets* y Modelos Ocultos de Markov

Informe de Proyecto de Graduación para optar por el título de Ingeniera en Electrónica con el grado académico de Licenciatura

María Gabriela Sáenz Meza

Cartago, Noviembre de 2006

INSTITUTO TECNOLOGICO DE COSTA RICA
ESCUELA DE INGENIERIA ELECTRONICA
PROYECTO DE GRADUACIÓN
TRIBUNAL EVALUADOR

Proyecto de Graduación defendido ante el presente Tribunal Evaluador como requisito para optar por el título de Ingeniera en Electrónica con el grado académico de Licenciatura, del Instituto Tecnológico de Costa Rica.

Miembros del Tribunal

Ing. Luis Paulino Méndez B, M.Sc

Profesor lector

Ing. Néstor Hernández, M.Sc

Profesor lector

Firma

Dr.-Ing. Pablo Alvarado Moya

Profesor asesor

Los miembros de este Tribunal dan fe de que el presente trabajo de graduación ha sido aprobado y cumple con las normas establecidas por la Escuela de Ingeniería Electrónica

Cartago, Costa Rica: noviembre de 2006

Declaro que el presente Proyecto de Graduación ha sido realizado enteramente por mi persona, utilizando y aplicando literatura referente al tema e introduciendo conocimientos propios.

En los casos en que he utilizado bibliografía, he procedido a indicar las fuentes mediante las respectivas citas bibliográficas.

En consecuencia, asumo la responsabilidad total por el trabajo de graduación realizado y por el contenido del correspondiente informe final.

Cartago, Noviembre de 2006

María Gabriela Sáenz Meza

Cédula: 1-1100-0396

Resumen

Como parte del proyecto de la “Red inalámbrica de telecomunicaciones para la protección del ambiente” se ha diseñado, implementado y comprobado por medio de procesamiento digital de señales, un prototipo algorítmico para el reconocimiento de patrones acústicos, específicamente de motosierras y disparos.

La meta del presente proyecto es desarrollar una herramienta para ayudar en la protección y conservación de los recursos naturales mediante la detección de actividades ilegales de tala y caza utilizando tecnología electrónica para evitar el uso irracional de dichos recursos.

En la implementación algorítmica de la etapa de extracción de características de la señal se utilizan *wavelets* y para el reconocimiento de la señal acústica los Modelos Ocultos de Markov (HMM).

Este documento describe las características técnicas y los principales resultados obtenidos, los cuales podrán ser utilizados de guía en la etapa posterior en la cual se integrará en hardware.

Palabras clave: Red inalámbrica, patrones acústicos, reconocimiento, *wavelets*, HMM, alfabeto.

Abstract

In the framework of the project "Wireless communication network for environmental protection" has been designed, implemented and verified an algorithmic prototype in order to recognize acoustic patterns, specifically chain saws and weapons.

The goal of the current project is to protect and conserve the natural resources, by means of illegal activities detection (felling of trees and hunts), using electronic technology to avoid their irrational use.

Wavelets are used in order to extract signal characteristics and Hidden Markov Models (HMM) are implemented to recognize acoustic patterns.

This document describes technical characteristics and the main results obtained, to serve as basis for the hardware implementation.

Key words: radio network, acoustic patterns, recognition, wavelets, HMM, codebooks.

Dedicatoria

A Luis Di, esto es para ti mi amorcito. Tú eres el motor para que mamá siga adelante y un ejemplo para mí, estoy orgullosa de ti. Gracias por toda tu paciencia en este tiempo.

A bebé, que viene en camino y lo estamos esperando con mucho amor.

A papi y mami, que durante toda mi vida han estado presentes apoyándome en las metas que quiero conseguir y en todas las pruebas que nos hemos encontrado en el camino. Gracias por todos los sacrificios, gracias por estar allí siempre con amor, que Diosito siempre los bendiga.

A Carlos, por creer en mí en todo momento, alentarme, apoyarme y estar conmigo en las buenas y en las malas. Mi amor, este esfuerzo por graduarme quiero imitarlo en el camino que vamos a emprender juntos.

A Laura y Rocío, por ayudarme con Luis Di, por sacrificar de su tiempo para que yo pudiera salir adelante en este proyecto.

A mis abuelitos que están pendientes de mis logros y que al igual que ellos siempre le piden a Dios que me acompañen y que todo me salga bien en la vida, yo también lo hago para que estén muchos años más a mi lado. Los quiero mucho!

Agradecimiento

A Dios por darme fuerzas para seguir adelante con todos mis proyectos.

A mi profesor Pablo, por compartir sus conocimientos que fueron de gran ayuda, además de sus consejos y recomendaciones durante todo el año, por su paciencia y solidaridad.

A los profesores Aris, Néstor y Luis Paulino por su apoyo y por estar atentos a mi desempeño.

A los profesores Alfonso y Gabriela, por sus buenos deseos y motivaciones para convertirme en ingeniera.

A todos los profesores de la Escuela de Ingeniería Electrónica del TEC que con su granito de arena me formaron durante este tiempo para ahora enfrentar de manera profesional la carrera que escogí.

INDICE GENERAL

Capítulo 1.	Introducción.....	1
1.1	Problema existente e importancia de la solución	1
1.2	Solución propuesta	3
1.2.1	Requerimientos	3
1.2.2	Descripción de la solución.....	3
Capítulo 2.	Meta y objetivos	7
2.1	Meta	7
2.2	Objetivo general	7
2.3	Objetivos específicos.....	7
Capítulo 3.	Marco teórico	8
3.1	Descripción del sistema	8
3.2	Antecedentes bibliográficos	12
3.3	Descripción de los principales principios físicos y/o electrónicos relacionados con la solución del problema	14
3.3.1	Etapa de extracción de características	14
3.3.2	Cuantificación de vectores (<i>Alfabetos</i>).....	23
3.3.3	K-means	25
3.3.4	Etapa de clasificación de la señal	25
Capítulo 4.	Procedimiento metodológico.....	34
4.1	Reconocimiento y definición del problema	34
4.2	Obtención y análisis de información	34
4.3	Evaluación de las alternativas y síntesis de la solución	35
4.4	Implementación de la solución	35
4.5	Reevaluación y rediseño	35
Capítulo 5.	Descripción detallada de la solución.....	37
5.1	Extracción de características.....	37
5.2	Definición de las características básicas de los Modelos Ocultos de Markov	40
5.3	Entrenamiento de los modelos	43
5.4	Reconocimiento de señales.....	44
Capítulo 6.	Análisis de resultados	46
6.1	Extracción de características de la señal	46
6.2	Definición de las características básicas de los Modelos Ocultos de Markov	52
6.3	Entrenamiento de los modelos	53
6.4	Evaluación	54
Capítulo 7.	Conclusiones y recomendaciones	59
7.1	Conclusiones	59
7.2	Recomendaciones	59
Bibliografía	61
Apéndices.....		65
A.1	Glosario, abreviaturas y simbología	65
A.2	Protocolos de medición	66
A.3	Organización de documentación en CD	72
A.4	Información sobre la institución.....	73

INDICE DE FIGURAS

Figura 1- 1 Diagrama de bloques de la solución.....	4
Figura 3- 1: Diagrama de bloques general de la “Red inalámbrica de telecomunicaciones para la protección del ambiente”	8
Figura 3- 2 Diagrama de flujo del sistema.....	9
Figura 3- 3 Diagrama de bloques de la solución para teledetectores.....	10
Figura 3- 4 Acondicionador de señal analógico	11
Figura 3- 5 Diagrama de bloques para análisis de filtros.....	17
Figura 3- 6 Bancos de filtros.....	18
Figura 3- 7 Representación del dominio de la frecuencia para la DWT	19
Figura 3- 8 Fourier versus wavelets [1]	21
Figura 3- 9 Diagrama de bloques para un alfabeto [38].....	24
Figura 3- 10 HMM para experimento de moneda	26
Figura 3- 11 HMM de primer orden	27
Figura 3- 12 Variable de forward α [15].....	31
Figura 3- 13 Variable de backward β [15]	32
Figura 6- 1 Señal acústica (a) Señal de longitud 2. (b) Señal de longitud 4	47
Figura 6- 2 Coeficientes de energía. (a) Señal de longitud 2. (b) Señal de longitud 4.....	49
Figura 6- 3 Vectores de energía promedio para el arma Carabina 22, 30 metros, 0 grados. (a) Señal de longitud 2. (b) Señal de longitud 4.....	50
Figura 6- 4 Distribución espacial de 5 dimensiones, ilustrado en 2 dimensiones.....	51
Figura 6- 5 Gráfico del cambio en porcentaje de acuerdo a la longitud de observaciones. Señales de entrenamiento longitud 2	57
Figura 6- 6 Gráfico del cambio en porcentaje de acuerdo a la longitud de observaciones. Señales de entrenamiento longitud 4	58

INDICE DE TABLAS

Tabla 6. 2 Variación de la longitud de observaciones entrenadas con señales de longitud 2.....	55
Tabla 6. 3 Variación de la longitud de observaciones evaluadas con señales de longitud 4.....	56

Capítulo 1. Introducción

1.1 Problema existente e importancia de la solución

En el ámbito internacional, Costa Rica es una oportunidad de éxito para implementar modelos de desarrollo sostenible debido a la importancia de los ecosistemas tropicales, su diversidad biológica y la capacidad de renovación de oxígeno, entre otras características [11]. Esta percepción que se tiene del país debe ser aprovechada para atraer la cooperación e inversión extranjera, recursos necesarios para la economía, los cuales se pueden dar, entre otros, mediante el turismo. El ambiente debe mantenerse equilibrado; se deben evitar la deforestación, principalmente la tala de árboles, y la caza ilegal de especies, muchas veces en peligro de extinción.

El proyecto “Red inalámbrica de telecomunicaciones para la protección del ambiente” [43], forma parte del esfuerzo que se realiza para frenar el deterioro acelerado de la naturaleza.

Durante el desarrollo de fases anteriores del proyecto se implementó la red de comunicación para el envío de datos a las casetas de los guarda-parques. Además se han diseñado dos prototipos para detectar motosierras y la implementación de la etapa de detección de fuego se trabaja en forma paralela e independiente a la etapa actual.

El prototipo diseñado anteriormente para detectar sonidos de motosierras cubre un radio de 150m [5,6]. Para la extracción de características de la señal se utilizaron filtros analógicos, los cuales constan de circuitos integrados y componentes electrónicos que demandarían mayor espacio en un prototipo que el diseño hecho de forma digital, además el consumo de potencia con el nuevo diseño de los circuitos podría minimizarse.

El clasificador empleado es muy simple por las limitaciones de la circuitería analógica y se consideran muy pocas características de la señal acústica para realizar la clasificación de la misma. Además el sistema no está entrenado de manera adecuada con los sonidos de interés y más comunes presentes en el lugar donde se desea implementar el prototipo, lo que aumenta las posibilidades de que el circuito falle en el momento de discriminar los sonidos.

Es importante mejorar los diseños propuestos anteriormente, especialmente en los siguientes aspectos:

- Que el prototipo algorítmico implementado en *hardware* sea de bajo consumo de potencia comparado con la exigida por un circuito analógico diseñado para el mismo objetivo.
- Aumentar el rango de detección del teledetector para señales acústicas definidas tanto de motosierras como de disparos.
- Optimizar la discriminación para evitar falsas alarmas u omisión de eventos importantes.
- Disminuir costos de implementación en cuanto a componentes electrónicos.
- Procurar que el prototipo sea más eficiente y eficaz en su accionar.

Se espera que al resolver el problema planteado electrónicamente surjan beneficios como:

- Prevenir desastres naturales.
- Evitar actividades ilegales de tala y caza.
- Conservar el ambiente
- Ofrecer a las autoridades una herramienta tecnológica que permita cubrir más áreas y en forma más oportuna su accionar en caso de que ocurran este tipo de actividades ilegales.

1.2 Solución propuesta

1.2.1 Requerimientos

Los requerimientos a los que está sujeto el desarrollo son dados por parte de los profesores encargados del proyecto en la Escuela de Ingeniería Electrónica:

- El análisis de la señal acústica se debe realizar por medios digitales.
- Los algoritmos que se desarrollen deben tener poca exigencia computacional. Esto llevará luego a mayor velocidad del sistema y bajo consumo de potencia cuando se implemente en *hardware*.
- Se puede realizar el análisis de la señal fuera de línea.
- Se debe tener un rango espacial considerable de acción para la detección de los sonidos de interés.
- La discriminación del circuito debe evitar falsas alarmas y no dejar pasar eventos importantes.

1.2.2 Descripción de la solución

El proyecto pretende desarrollar una herramienta que permita detectar actividades ilegales como la caza y tala en zonas protegidas de forma instantánea, diseñando un prototipo algorítmico que utilice el procesamiento digital para determinar por análisis de señales acústicas la presencia de motosierras o disparos y cumpla con las especificaciones dadas. Esto alertaría de forma oportuna a las autoridades respectivas colaborando con la protección del ambiente.

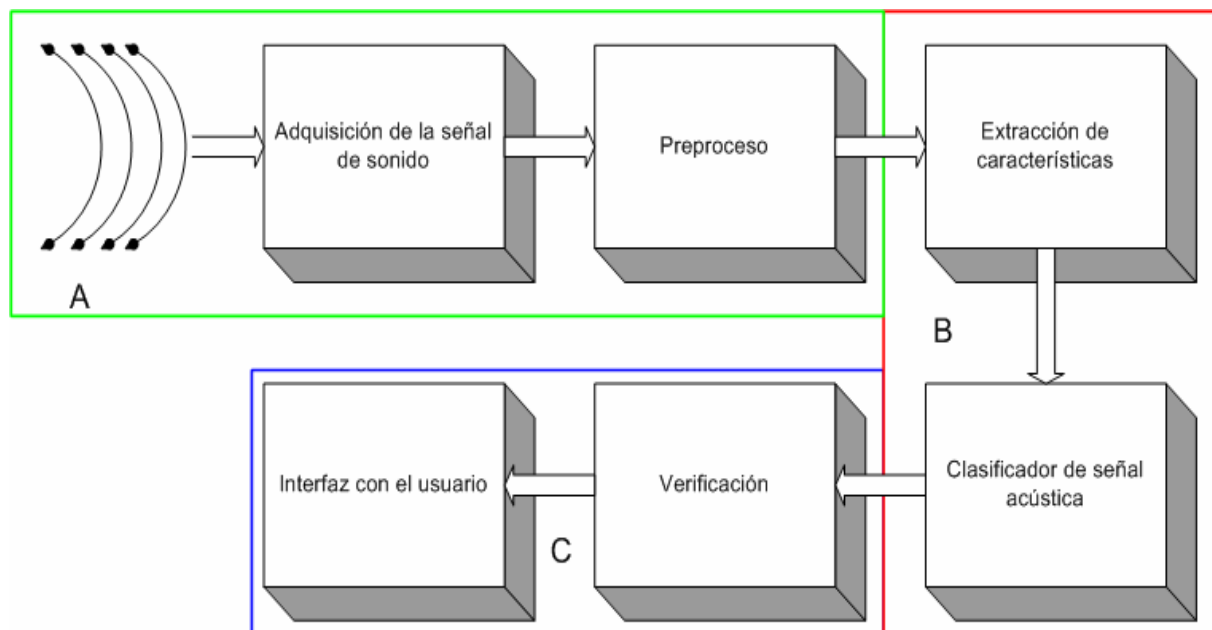


Figura 1- 1 Diagrama de bloques de la solución

El prototipo algorítmico se realiza utilizando MatLab, donde es necesario adquirir previamente las señales acústicas de interés, lo que se implementa en la etapa A del diagrama mostrado en la figura 1-1.

Para ello se estableció una base de datos con sonidos específicos de motosierras y armas de fuego utilizadas normalmente en las zonas boscosas de nuestro país. Además se grabaron sonidos comunes del ambiente natural (lluvia, grillos, truenos) en sitios donde se encontrarán ubicados posiblemente los teledetectores, junto con otros sonidos ajenos al bosque como avionetas o camiones, los cuales podrían interferir en la señal de interés (ver apéndice A.3).

Estos archivos de sonido no solo servirán para la experimentación actual, sino que en las siguientes investigaciones se tendrá la manera de comprobar la funcionalidad del prototipo, ya que se compararían los resultados obtenidos con la misma base de datos.

En la etapa A diseñada se acondiciona la señal mediante software para que se normalice la amplitud del sonido que será la entrada luego a la etapa B (ver figura 1-1).

La etapa B es la principal en el desarrollo del proyecto, ya que en ésta se reconocen digitalmente patrones de sonidos de motosierras y disparos, utilizando la teoría de procesamiento digital de señales (DSP) primordialmente.

En el módulo de extracción de características se definen las particularidades para clasificar la señal en proceso (frecuencia de muestreo y bandas de frecuencia, ancho de palabra o número de muestras presentes, tipo de adquisición). Determinado ese punto, entonces las componentes del sonido en que se basa la clasificación de si es o no efectivamente un disparo o una motosierra se desarrollan mediante la teoría de *wavelets*.

Las *wavelets* son funciones localizables en el tiempo y en frecuencia, lo que permite saber el contenido en frecuencia de la señal acústica y cuándo éstas componentes de frecuencia se encuentran presentes en la señal.

Este módulo selecciona las componentes de frecuencia de la señal de forma digital en las que se basa la clasificación entre el sonido del disparo, la motosierra y todos los demás sonidos.

Luego, para la clasificación de los sonidos se propone utilizar como clasificadores estadísticos los Modelos Ocultos de Markov (*HMM*) debido al uso exitoso que ha tenido en diferentes campos como el reconocimiento de voz y predicción de fallas entre otros [16,17,26,27].

El sistema se entrena con disparos y sonidos de motosierra a diferentes distancias, generando diferentes clases (a 30 m, 90 m, 150m y 600m), y las distancias

anteriores a diferentes ángulos con respecto la dirección del sonido de interés (0, 90 y 180 grados). El rango del sensor dependerá entonces del cuerpo de entrenamiento donde hay que tomar en cuenta las características del bosque. Mientras la señal acústica sea perceptible al micrófono utilizado, se puede entrenar al prototipo algorítmico.

La forma de verificar los resultados es con la experimentación iterativa, ya que no se cuenta con un antecedente directo de reconocimiento digital de sonidos motosierras y disparos para la protección del ambiente.

Las interfaces gráficas de usuario utilizadas en la etapa C serán las necesarias para la experimentación del proyecto, aún no se implementará una para el uso de personas que vayan a utilizar la aplicación final.

El establecimiento específico de los teledetectores, la cantidad y la forma adecuada en el lugar se decidirá cuando se experimente con los mismos, dependiendo de los obstáculos presentes que afecten la transmisión de la señal acústica y los fenómenos que éstos puedan generar (refracción, reflexión, difracción, dispersión, disipación).

Capítulo 2. Meta y objetivos

2.1 Meta

- Proteger y conservar los recursos naturales mediante la detección de actividades ilegales de tala y caza utilizando tecnología electrónica para evitar el uso irracional de los mismos.

2.2 Objetivo general

- Diseñar e implementar por medios digitales un prototipo algorítmico para el reconocimiento de sonidos de motosierras y disparos en el sistema de red inalámbrica de telecomunicaciones con el fin de proteger el ambiente.

2.3 Objetivos específicos

- Establecer una base de datos con los sonidos conocidos de motosierras y armas de fuego específicas y sonidos clásicos de las zonas boscosas del país para la verificación experimental de los algoritmos a diseñar.
- Extraer características relevantes de la señal original de sonido utilizando la teoría de *wavelet*.
- Evaluar mediante *HMM* la clasificación correcta de la señal en estudio.
- Redactar un documento final que describa las características técnicas y los principales resultados obtenidos para que se utilice de guía en el proyecto posterior integración en *hardware*.

Capítulo 3. Marco teórico

3.1 Descripción del sistema

El proyecto “Red inalámbrica de telecomunicaciones para la protección del ambiente”, pretende frenar el la destrucción acelerada de la naturaleza utilizando la ingeniería electrónica como herramienta y con ello ayudar al personal encargado de vigilar nuestros bosques a que la conservación sea mucho más eficiente y eficaz.

El proyecto consta de los siguientes módulos:

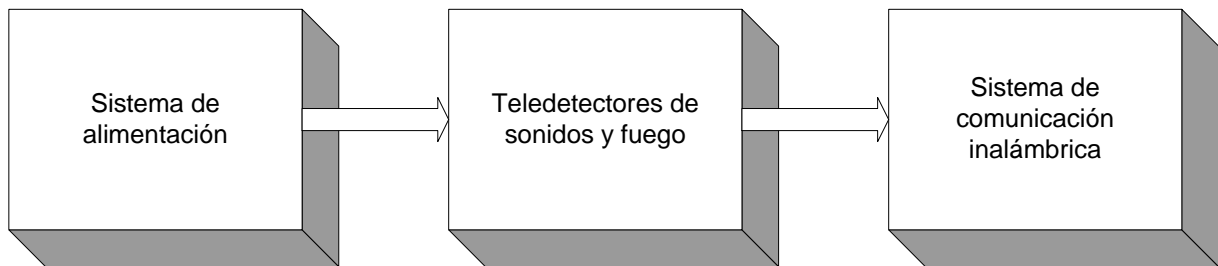


Figura 3- 1: Diagrama de bloques general de la “Red inalámbrica de telecomunicaciones para la protección del ambiente”

Se debe considerar la forma de alimentar el sistema para especificar los requerimientos del consumo de potencia, lo cual se determinará cuando se obtenga el prototipo final.

El proceso que debe realizar el sistema se muestra en la figura 3-2:

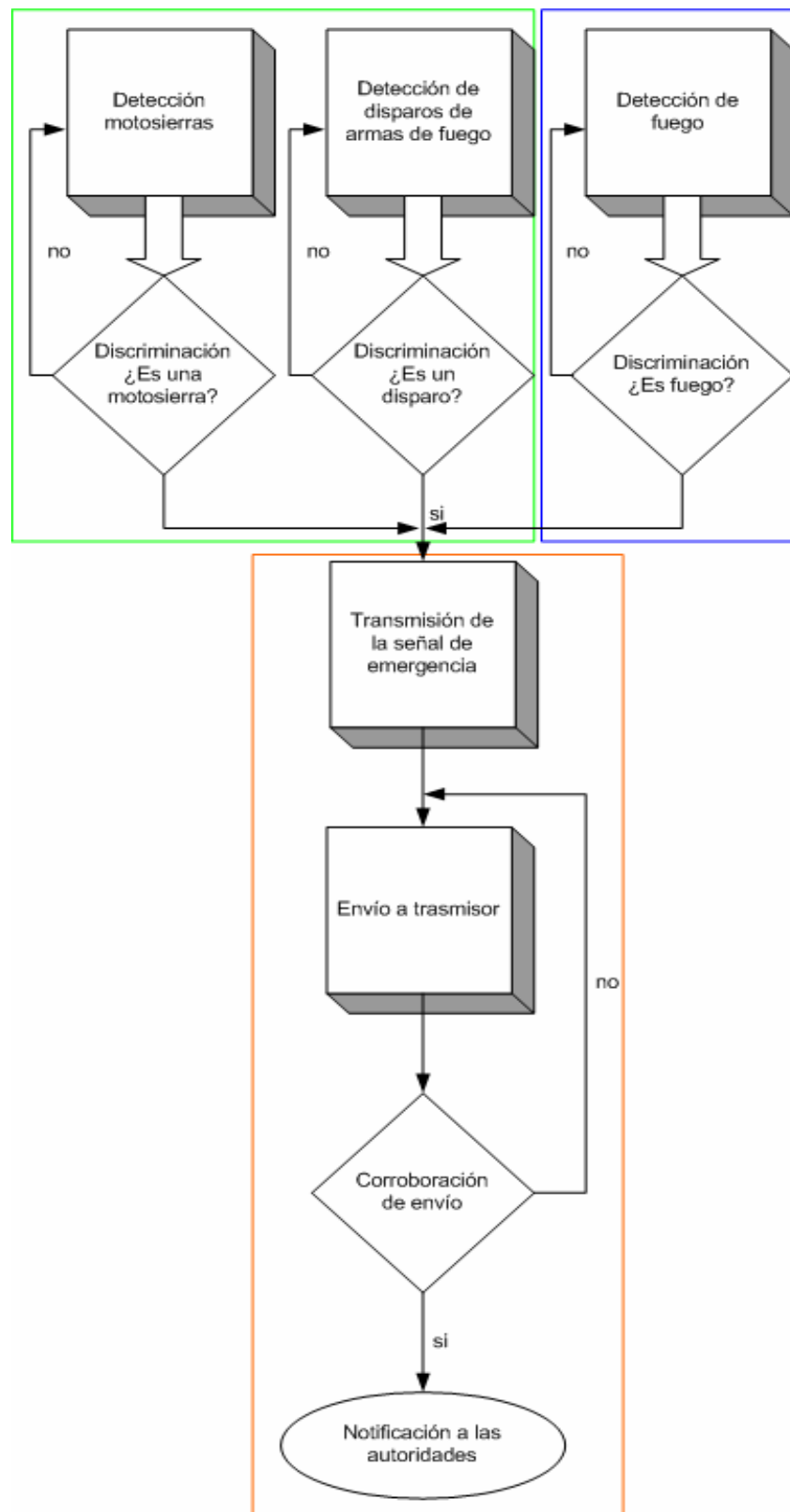


Figura 3- 2 Diagrama de flujo del sistema

Estos módulos forman una red de sensores teledetectores que podrían cubrir gran parte de las zonas que se quieren proteger ya sea de la tala y caza ilegal e inclusive del fuego. Se detecta la señal, ésta se clasifica para saber de que tipo es y si corresponde a las antes mencionadas se activa la emergencia que es enviada a las autoridades más cercanas al lugar del hecho.

En la figura 3.3 se presenta el proceso mediante el cual la señal acústica será clasificada según sea el sonido de interés (disparo o motosierra) y sus etapas son explicadas a continuación:

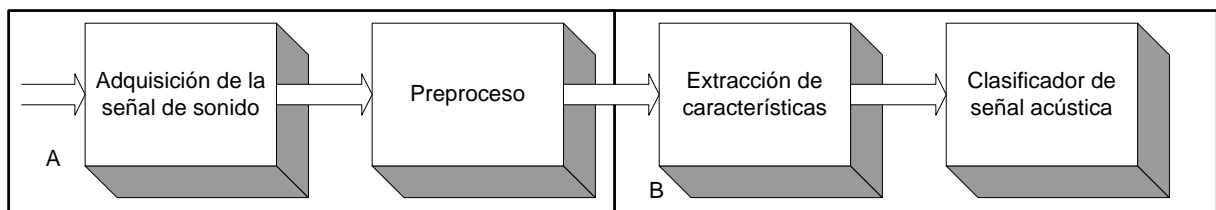


Figura 3- 3 Diagrama de bloques de la solución para teledetectores

Etapas A:

El módulo de adquisición utilizado en el prototipo analógico anterior es presentado en [5,6]. Éste módulo se encargó de transformar la señal a eléctrica y consta de un micrófono (WM-64MN), el cual es omnidireccional por lo que permite cubrir mayor área con un solo micrófono.

El micrófono actúa como filtro paso bandas ya que tiene un rango de frecuencias limitado de los 20 Hz a los 16 kHz y es apto para soportar el clima del bosque. Esta etapa fue amplificada a su salida ya que los voltajes entregados del micrófono por las variaciones de capacitancia (es un micrófono de condensador) son bastante bajos.

Para acoplar la baja impedancia del micrófono con la impedancia de entrada del acondicionador de señal (pre-proceso), se utilizaron operacionales TL072 con sus

respectivas ganancias y para amplificar aún más el voltaje de salida un circuito integrado NE5532 en configuración inversora realiza esta tarea. [5]

El módulo de preproceso acondicionó la señal a un rango de voltaje regulado (intensidad del sonido) mediante normalización, lo que permite que el resto del sistema trabaje con los mismos niveles de referencia en cuanto a amplitud, como lo indica la siguiente ecuación:

$$V' = \frac{V_{entrada}}{V_{promedio} + k} \quad (3.1)$$

Esta etapa fue configurada analógicamente de la siguiente manera:

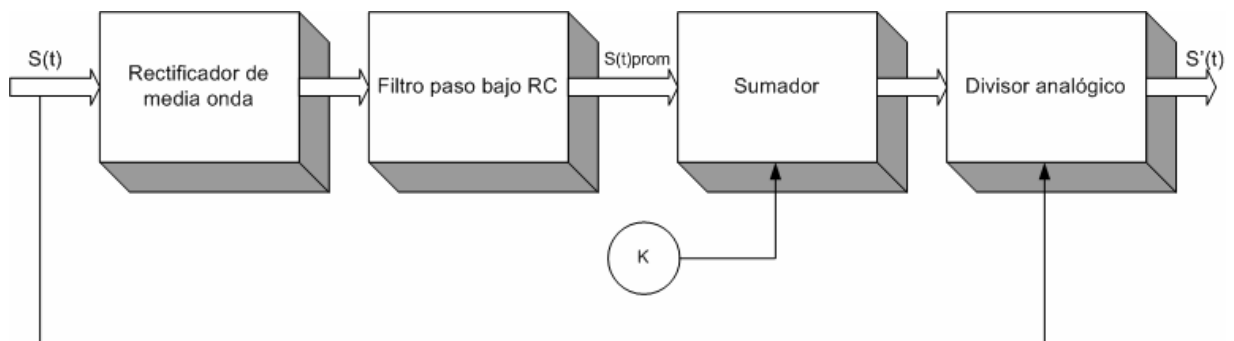


Figura 3- 4 Acondicionador de señal analógico

Etapas B:

El teledetector en la fase anterior del proyecto se diseñó utilizando técnicas de reconocimiento acústico basado de clasificación estadística de patrones implementados mediante un banco de filtros y un clasificador lineal simple. Este clasificador utilizó una red neuronal de una capa y permitió discriminar diferentes sonidos.

3.2 Antecedentes bibliográficos

En el área de “Reconocimiento de patrones acústicos”, se identifican esquemas de sonido en diferentes contextos, por ejemplo en el ambiente de aeropuertos, alarmas, voz humana, entre otros. En zonas boscosas para evitar la caza y la tala no hay referencia.

En [16] se presenta una unidad acústica de la vigilancia (ASU) construida en el marco de un proyecto para la localización de fuentes audio. La unidad se hace para localizar los vehículos que emiten sonidos en la gama de [10Hz-300Hz] con una exactitud de un grado y una distancia aproximada de 300m.

El reconocimiento automático de lenguaje (ASR) analizado en publicaciones de la IEEE (*Speech and Audio Processing, Computer Speech and Language, and Speech Communications*) fue basado en la clasificación de los patrones acústicos extraídos de la señal de un discurso [3].

Además [44] describe un sistema que divide señales en segmentos musicales según la presencia o la ausencia del instrumento.

La empresa SINTEF ha tratado el problema de reconocimiento de patrones acústicos de diferentes campos: En colaboración con una empresa determinada de aviación civil (Avinor) y Norsonic, desarrolló un sistema de supervisión que registra todo el ruido cerca de un aeropuerto, y extrae el ruido que es debido al tráfico aéreo solamente. Esta clasificación se basa en redes neuronales [9].

Dentro del campo de la salud en [9] se trabajó con un sistema que apoya el diagnóstico relacionado con roncar, analizando dicho sonido y localizando las posiciones corporales donde ocurre la generación de los mismos.

Las explosiones impulsivas de la energía acústica son una fuente rica de la información en el mundo natural. Los sistemas biológicos los procesan rápida y económicamente [4]. Se describe un circuito analógico biológico con arquitectura de integración (VLSI) para la clasificación en tiempo real de señales acústicas. La normalización de las señales de tiempo-frecuencia permite la puesta en práctica de un algoritmo de correlación. El algoritmo substituye la multiplicación analógica-analógica por la multiplexación binaria de señales analógicas, esto elimina la necesidad del almacenaje y de la multiplicación analógica. Este desarrollo se basa en la manera del proceso acústico inteligente en usos de baja potencia tales como teléfonos portátiles.

En la bibliografía citada se tiene en común el uso de redes neuronales, Modelos Ocultos de Markov, clasificadores estadísticos y el procesamiento de la señal por medios digitales dependiendo de la aplicación.

Los Modelos Ocultos de Markov se presentan en [15,22-24], donde claramente se explica la estructura matemática y la base teórica para el uso en una amplia gama de campos. Cuando los modelos son aplicados correctamente, trabajan muy bien en la práctica. En publicaciones de la IEEE se procura repasar cuidadosa y metódicamente los aspectos teóricos de este tipo de modelo y de la demostración estadística de cómo se han aplicado en problemas seleccionados de reconocimiento.

De [25-28] además de aplicar HMM, se estudia la caracterización de las señales tanto con análisis de Fourier como con transformadas *wavelets*, para producir la secuencia de los vectores espectrales que pueden capturar exactamente las características salientes de dichas señales. Conceptos como la cuantificación del vector se utilizan para asignar la secuencia de observaciones de la fase de entrenamiento. Durante la etapa de prueba, la secuencia se compara contra todos los modelos ocultos desarrollados de Markov y se clasifica.

3.3 Descripción de los principales principios físicos y/o electrónicos relacionados con la solución del problema

La teoría tanto de *wavelets*, alfabeto, *K-means* y Modelos Ocultos de Markov será presentada a continuación, definiendo lo más relevante para dichos temas. Serán desarrollados algunos conceptos matemáticos que son aplicados posteriormente en la solución.

Parte de los temas de análisis de Fourier, redes neuronales artificiales, y clasificadores estadísticos han sido previamente investigados para la red de telecomunicación inalámbrica para la protección del ambiente [5,6], como base para la toma de decisiones en cuanto a la solución seleccionada.

3.3.1 Etapa de extracción de características

Diferentes procesos del mundo real producen generalmente salidas observables las cuales se pueden caracterizar como señales [15]. Dichas señales pueden ser discretas (como los vectores cuantificados de un *alfabeto*), o continuas. La fuente de la señal puede tener características estadísticas que no varían con el tiempo, o no estacionaria donde, las características de la señal varían en un cierto plazo. Las señales pueden ser puras (de una sola fuente), o se podrían corromper con otras fuentes de la señal como el ruido o distorsiones de la transmisión, reverberación, etc. La extracción de características de las señales en estudio se hará basado en *wavelets*, dichas señales son no estacionarias y podrían tener presente ruido.

Fundamento teórico de wavelets

Matemáticamente los *wavelets* y su análisis, se refieren a la representación de una señal en términos de la combinación aditiva de otras señales generadas de traslaciones y contracciones de otra señal de longitud finita o con forma de onda

oscilante que decae rápidamente, la cual se conoce como *wavelet* madre o fundamental [1].

La palabra *wavelet* es debida a Morlet y Grossman a inicios de los años 80, donde utilizaron *ondelette* derivado del francés, el cual significa pequeña onda. Tiempo después *onde* fue traducido al inglés como *wave*, resultando finalmente *wavelet*.

Las transformadas *wavelets* son clasificadas ampliamente en la transformada discreta de *wavelet* (DWT) y la transformada continua de *wavelet* (CWT). La principal diferencia entre ambas es que la transformada continua opera sobre cada escala posible de la señal, mientras que la discreta usa un subconjunto específico de todos los valores de dicha escala. Detalles de estas transformadas se pueden ver en [35].

CWT

En el procesamiento de la señal, la CWT de una función f , esta definida por:

$$\gamma(\tau, s) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{|s|}} \overline{\psi\left(\frac{t-\tau}{s}\right)} dt \quad (3.2)$$

Donde τ representa la traslación, s la escala y ψ la *wavelet* madre. En la transformada continua de *wavelet*, una señal dada de energía finita se proyecta en una familia continua de las bandas de frecuencia, por ejemplo en cada banda de frecuencia de la forma $[f, 2f]$ para todas las frecuencias positivas $f > 0$. La integración de los componentes de frecuencia puede reconstruir la señal original.

Las bandas de frecuencia o subespacios son versiones escaladas de un subespacio de escala 1. Este subespacio alternadamente está en la mayoría de las situaciones generadas por los cambios de una función generada: la *wavelet* madre.

$$\psi \in L^2(\mathfrak{R}) \quad (3.3)$$

En un subespacio de escala a o con una banda de frecuencia de $[1/a, 2/a]$ es generada por las funciones conocidas como *wavelet* hijas:

$$\psi_{a,b}(t) = \frac{1}{a} \psi\left(\frac{t-a}{b}\right) \quad (3.4)$$

Donde a es positivo y define la escala y b es cualquier número real que define el cambio. Además la proyección de una función x sobre un subespacio de escala a tiene la forma:

$$x_a(t) = \int_{\mathfrak{R}} WT_{\phi}\{x\}(a,b) * \psi_{a,b}(t) db \quad (3.5)$$

Con coeficientes *wavelets* dados por:

$$WT_{\phi}\{x\}(a,b) = \langle x, \psi_{a,b} \rangle = \int_{\mathfrak{R}} x(t) \psi_{a,b}(t) dt \quad (3.6)$$

DWT

Este tipo de transformada se refiere a los *wavelets* que son muestreados discretamente. Dentro de las características generales de una transformada *wavelet* se tiene en primera instancia que puede ser desarrollada en $O(n)$ operaciones, y en segundo lugar captura no solo alguna notación del contenido de la frecuencia de entrada examinándola en diferentes escalas sino también el contenido temporal (el

tiempo en el cual esta frecuencia ocurre). Combinando las dos propiedades anteriores se obtiene la transformada rápida de *wavelet* (FWT), una alternativa de la convencional transformada rápida de Fourier.

El conjunto más común de transformadas *wavelet* está basado en el uso de las relaciones de repetición para generar muestras discretas progresivamente más finas de una función implícita de la *wavelet* madre, donde cada resolución es dos veces el de la escala anterior.

La DWT de una señal x es calculada por una serie de filtros, donde en el primer paso un filtro paso bajo con respuesta al impulso g es aplicado a las muestras [1]:

$$y[n] = (x * g)[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot g[n-k] \quad (3.7)$$

La señal también es descompuesta simultáneamente usando un filtro paso alto h . Con ambos filtros las salidas resultantes son los coeficientes de detalle dados por el filtro paso alto y los coeficientes de aproximación, dados por el filtro paso bajo, según la figura 3-5:

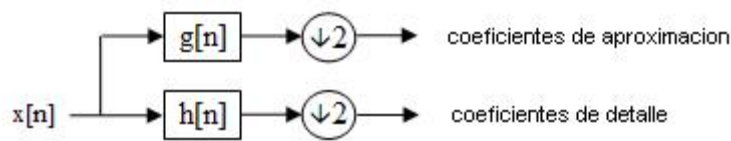


Figura 3- 5 Diagrama de bloques para análisis de filtros

Ambos filtros se representan de la siguiente manera:

$$y_{bajo}[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot g[2 \cdot n - k] \quad (3.8)$$

$$y_{alto}[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot h[2 \cdot n - k] \quad (3.9)$$

Esta descomposición ha dividido en dos la resolución del tiempo puesto que solamente la mitad de cada salida del filtro caracteriza la señal, pero se tienen en la salida dos bandas de frecuencia, así que se ha doblado la resolución de la frecuencia.

La relación entre ambos filtros con una longitud de L es:

$$h[L-1-n] = (-1)^n * g[n] \quad (3.9)$$

La resolución se mejora repitiendo este proceso para la salida y_{bajo} , utilizando un banco de filtros, donde cada nivel de la señal es descompuesto en bajas y altas frecuencias. Lo anterior es representado como un árbol binario con los nodos más significativos con diversa localización de tiempo-frecuencia como es mostrado en la figura 3-6:

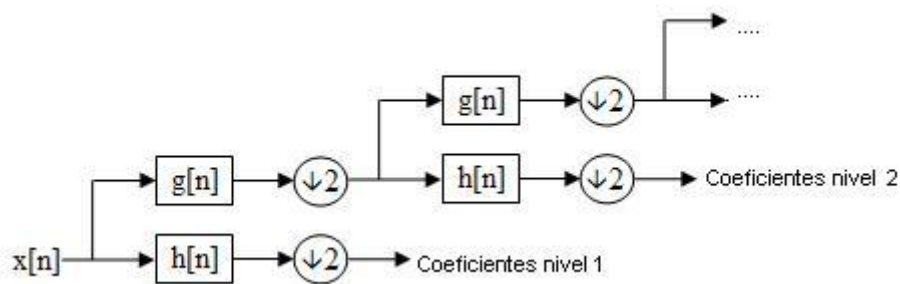


Figura 3- 6 Bancos de filtros

Debido a este proceso de descomposición de la señal de entrada debería ser un múltiplo de 2^n , donde n es el número de niveles [1]. Por ejemplo una señal con $N=16$ muestras, un rango de frecuencia de 0 a f_n y 3 niveles de descomposición, 4 escalas de salida son producidas de la siguiente manera, lo cual refleja que es $2 \cdot \log N$ (ver figura 3-7):

Nivel de f	Frecuencia	# de Muestras
3	0 a $f_n/8$	4
	$f_n/8$ a $f_n/4$	4
2	$f_n/4$ a $f_n/2$	8
1	$f_n/2$ a f_n	16

Tabla 3- 1: Niveles de frecuencia

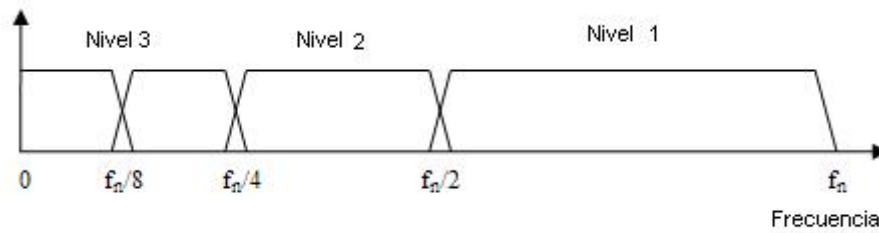


Figura 3- 7 Representación del dominio de la frecuencia para la DWT

La base de los *wavelets* construida por Daubechies [32] es de las técnicas más populares para el análisis de la señal. Los parámetros de la descomposición de la transformada discreta de *wavelets* están dados por una secuencia de valores reales ($h[0], h[1], \dots, h[n]$) que soluciona las ecuaciones de cuadratura:

$$\sum_{n=0}^N h[n] = \sqrt{2} \quad (3.10)$$

$$\sum_{n=0}^N (-1)^n n^k h[n] = 0 \quad (3.11)$$

$$\sum_{n=0}^{N-2k} h[n] h[n+2k] = 0 \quad (3.12)$$

Donde (3.11) tiene una $k=0,1,2,\dots, \frac{(n-1)}{2}$ y (3.12) una $k=1,2,\dots, \frac{(n-1)}{2}$.

Para el desarrollo matemático detallado de las ecuaciones anteriores se puede referir a [32], otra fuente de cómo extraer las características de la señal utilizando *wavelets* esta dado por Wang en [37].

Para extraer las características de la señal discreta *wavelet*, cada coeficiente de un vector $\mathbf{E}=[E_1, E_2, \dots, E_J]$ puede ser representado por el promedio de su energía, la cual esta definida por:

$$E_j = \frac{1}{T_j} \sum_{k=1}^{T_j} g_j^2(k) \quad (3.13)$$

Donde T_j es el número de coeficientes en cada escala. El valor de T_j es diferente para diversas escalas, porque el espaciamiento del tiempo de los coeficientes *wavelet* varía con la resolución j . Como resultado, las observaciones son independientes de las magnitudes de la señal. Cada característica es un vector único que corresponde a algunas características de los estados del sistema diseñado.

Análisis de Fourier vs Wavelet:

El término *wavelet* significa “pequeña onda” o función localizable en el tiempo, que visto desde una perspectiva del análisis o procesamiento de señal puede ser considerada como una herramienta matemática para la representación y segmentación de señales, análisis tiempo - frecuencia y fácil implementación de rápidos algoritmos computacionales [35].

En cuanto a sus aplicaciones, la transformada *wavelet* discreta se utiliza para la codificación de señales, mientras la continua se aplica en el análisis de las mismas, a menudo sustituyendo la transformada de Fourier.

En términos históricos, el desarrollo de las *wavelets* se da a partir del trabajo de Alfred Haar a principios del siglo XX [1]. Todas las transformaciones *wavelet* pueden ser consideradas formas de representación en tiempo-frecuencia y por tanto, están relacionadas con el análisis armónico.

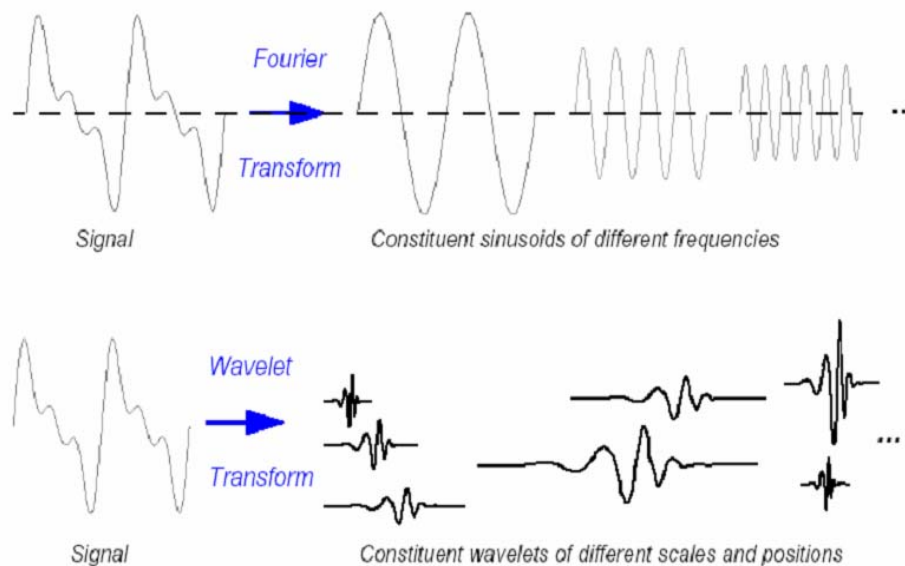


Figura 3- 8 Fourier versus wavelets [1]

De la figura 3-8 se observa que el análisis de *wavelets* está especialmente indicado para señales con pulsos, intermitencias o sucesos que ocurren de manera no periódica. Para estas señales, el análisis de Fourier da muy poca información, al perder toda información temporal.

La transformada *wavelet* representa a una señal en términos de una forma de onda de longitud finita que decae con rapidez: *wavelet* madre. Esta forma de onda es escalada y trasladada para corresponder con la señal de entrada.

La transformada de Fourier constituye una herramienta mediante la cual se puede obtener información sobre cómo está distribuida la energía de una señal en sus distintas componentes de frecuencia, es decir, se pueden conocer todas las

componentes de frecuencia existentes en la señal y sus respectivos aportes energéticos. Todo lo anterior se puede resumir diciendo que la transformada de Fourier tiene una perfecta resolución en frecuencia lo que la hace una herramienta muy útil para el análisis de señales estacionarias. Sin embargo, ella no puede ser aplicada con el objeto de obtener información precisa de cuándo o dónde las diferentes componentes de frecuencia se encuentran en la señal como es el caso no estacionario cuyo contenido espectral varía con el tiempo.

El análisis de Fourier es inestable frente a señales no estacionarias, donde al añadir un impulso localizado en el tiempo a una señal, todo el espectro de Fourier se ve afectado, mientras que solo algunos coeficientes de las *wavelets* se modificarán.

Cuando un sistema es lineal todos los modos de vibración son modos propios del sistema y el análisis de Fourier describe completamente la relación entrada/salida de los mismos, pero si no es así, la descomposición en modos propios no da información interesante, ya que se mezclan modos de oscilación.

La transformada discreta de *wavelet* presenta además claras ventajas frente a la transformada de Fourier [35]:

- Más rápida desde el punto de vista computacional.
- En muchos casos proporciona un mejor ajuste a los datos con menos coeficientes (permitiendo una mejor compresión de los datos que los métodos basados en Fourier).
- Las técnicas de filtrado de ruido basadas en *wavelet* dan mejores resultados.

Estas ventajas son de interés en este proyecto ya que cumplen con los requerimientos propuestos en el capítulo 1.

Aunque en los últimos años se ha hecho un gran esfuerzo por darle a las transformadas *wavelets* todo el rigor matemático que tiene la transformada de Fourier y unificar métodos y notaciones, el ritmo de aparición de publicaciones sobre el tema hace que no sea tarea fácil [32]. Además no permite realizar algunos cálculos como los relacionados con la convolución o la modulación de una señal.

3.3.2 Cuantificación de vectores (*Alfabetos*)

La cuantificación de vectores es usada en muchas aplicaciones como la compresión de imagen y sonido, reconocimiento de voz y en general en patrones de reconocimiento. [38].

Un vector cuantificado mapea k-dimensiones de vectores en un espacio R^k dentro de un conjunto finito de vectores $Y = \{y_i : i=1,2,\dots,M\}$. Cada vector y_i es llamado símbolo y el conjunto de todos los símbolos es llamado alfabeto. Asociado con cada vector y_i esta una región muy cercana llamada región de Varonoi y es definida por:

$$V_i = \{x \in R^k : \|x - y_i\| \leq \|x - y_j\|, \text{ para todo } j \neq i\} \quad (3.14)$$

El símbolo representativo de un vector de entrada es aquel más cercano con respecto a cada vector compuesto por el alfabeto según la distancia euclidiana. La distancia euclidiana se define por:

$$d(x, y_i) = \sqrt{\sum_{j=1}^k (x_j - y_{ij})^2} \quad (3.15)$$

Donde x_j es el vector de entrada de la señal a evaluar y y_{ij} es cada vector del alfabeto.

Un vector cuantificado está compuesto por dos operaciones, la primera es codificar y la segunda decodificar. El codificador toma la entrada del vector y como salida se obtiene el índice del *símbolo* que resulta con la distorsión más baja.

En este caso la distorsión más baja es encontrada evaluando la distancia euclidiana dada en (3.15) entre el vector de la entrada y cada *símbolo* en el *alfabeto*. Una vez que se encuentre el *símbolo* más cercano, el índice de ese *símbolo* se envía a través de un canal (podría ser una memoria interna, un canal de comunicaciones, y así sucesivamente). Cuando el decodificador recibe el índice del *símbolo*, sustituye el índice por el *símbolo* asociado. El diagrama de bloques para este proceso se presenta en la figura 3-9:

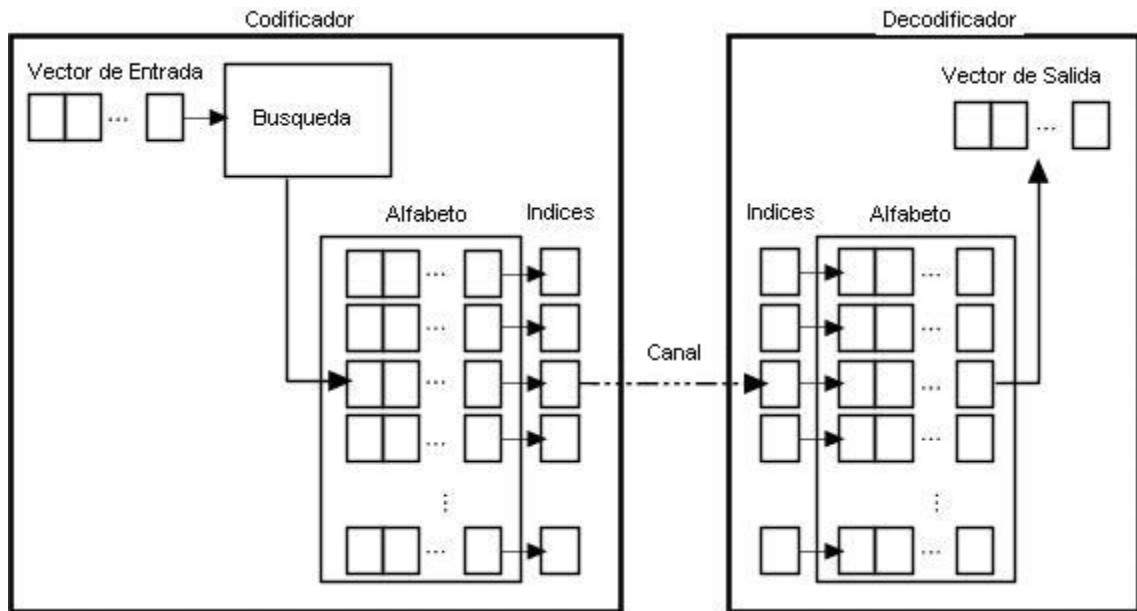


Figura 3- 9 Diagrama de bloques para un alfabeto [38]

El índice I^* del vector cuantificado esta explicado por [39]:

$$I^* = \arg \min \left[d \left(x - \overline{y_i} \right) \right] \quad (3.16)$$

3.3.3 K-means

El algoritmo de *K-means* será el utilizado para generar el diseño del *alfabeto* en el proyecto.

El resultado de un “*clustering*” se puede utilizar para una representación compacta de un sistema de objetos. Este método se refiere a la cuantificación del vector. El método de LBG para el diseño de cuantificación del vector es una generalización del algoritmo de *K-means*, éste se realiza en los objetos representados en una característica del espacio.

Primero, se seleccionan los centroides iniciales, después usando la función de la distancia mínima, todos los objetos se clasifican en los *clusters* apropiados. Una medida de la distorsión para la clasificación actual, (ej: error de la media cuadrática) se calcula y cada centro del *cluster* se actualiza para ser el valor medio de los vectores característicos que corresponden a los objetos dentro del mismo. En esta etapa, se reclasifican los objetos, se calculan los nuevos centros, y la medida de la distorsión es actualizada. El proceso de la clasificación y de la actualización se realiza hasta normalizar el cambio en distorsión.

Los detalles de su implementación se desarrollarán en el capítulo de la solución al problema y también se puede referir a [38 y 42].

3.3.4 Etapa de clasificación de la señal

Modelos Ocultos de Markov (HMM)

Un modelo oculto de Markov (*HMM*) es un modelo estadístico para el estudio de señales discretas o continuas, escalares o vectoriales, con datos puros o corruptos.

Para el entendimiento del concepto se presenta el ejemplo del modelo donde se tira una moneda y no se sabe si el resultado es “escudo o corona” [15]:

Se asume que se está en un cuarto con una barrera que no permite la visibilidad de lo que pasa al otro lado, por ejemplo una cortina, y al otro lado de la cortina hay otra persona que tira una o varias monedas hacia arriba. La otra persona no dice lo que está haciendo exactamente, solo el resultado de cada vez que tira la moneda. Así se realiza una secuencia del experimento de tirar la moneda oculta, con una secuencia de observaciones que consiste de una serie de caras y cruces, por ejemplo una observación típica de la secuencia podría ser: $O = O_1, O_2, \dots, O_T = E E C C C E C C E \dots E$, donde C es para corona y E para escudo.

Dado el panorama, el problema de interés es como construir un HMM para explicar (modelo) la secuencia observada de escudos y coronas. El primero de los problemas es decidir qué estados del modelo corresponden a qué y después decidir cuántos estados podría haber en el modelo. Una posible opción es decidir que solo una moneda esta siendo tirada, en este caso se podría modelar la situación con dos estados, donde cada estado corresponde a un lado de la moneda. Este modelo se puede observar en la figura 3-10:

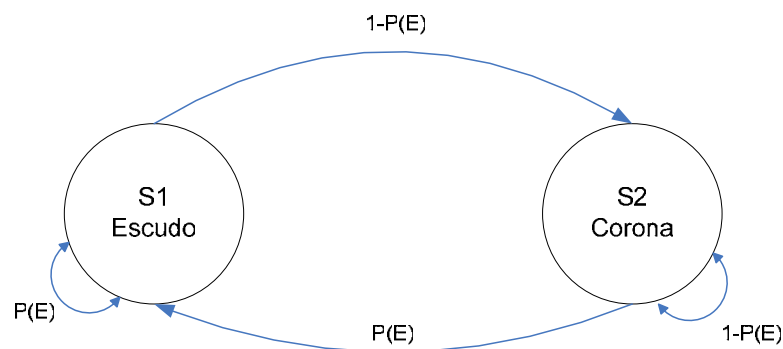


Figura 3- 10 HMM para experimento de moneda

En este caso el modelo de Markov es observable y la única edición para la especificación completa del modelo sería decidir sobre el mejor valor de probabilidad que tiene para caer la moneda la próxima vez que se tire.

Este concepto matemático se deriva de las cadenas de Markov [33], el cual se explica a continuación.

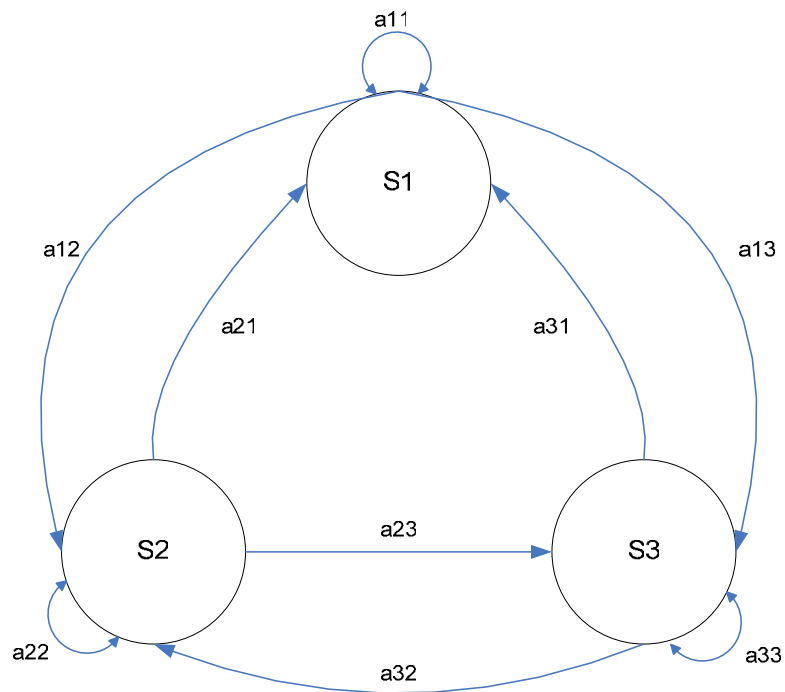


Figura 3- 11 HMM de primer orden

El sistema de la figura 3-11 puede iniciar en alguno de los N estados (S_1, S_2, S_3), en cualquier instante del tiempo t . El correspondiente Modelo Oculto de Markov para las observaciones discretas se caracteriza por:

- El número $N=(S_1, S_2, S_3, \dots, S_N)$ de estados ocultos del modelo. Estos estados son interconectados de tal manera que cualquier estado oculto se pueda alcanzar desde cualquier otro estado oculto.

- El número M de distintas observaciones o símbolos observables que se determina con el procedimiento del alfabeto visto en la sección 3.3.3 y se denota como $\mathbf{V}=\{v_1, v_2, v_3, \dots, v_n\}$.
- La distribución de probabilidad de transición entre estados $A=[a_{ij}]$ se denota por:

$$a_{ij} = P(q_t = S_j | q_{t-1} = S_i), \quad 1 \leq i, j \leq N \quad (3.17)$$

- La distribución de probabilidad de observaciones $B=[b_{jm}]$ define la distribución en el estado j y O_t es la observación en el tiempo t .

$$b_{jm} = P(o_t = v_m | q_t = S_j), \quad 1 \leq m \leq M \quad (3.18)$$

- Finalmente la distribución del estado inicial esta dada por $\pi=[\pi_1]$ definiéndose como:

$$\pi_i = P(q_i = S_i), \quad 1 \leq i \leq N \quad (3.19)$$

La especificación completa de un HMM requiere tres conjuntos de medidas probabilísticas: A , B y π , utilizándose la notación: $\lambda=(A, B, \pi)$.

Es llamada clase a las salidas esperadas del modelo y cada una de ellas debe ser entrenada por un HMM distinto.

Hay 3 problemas principales a solucionar con los Modelos Ocultos de Markov:

- Dados los parámetros del modelo, calcular la probabilidad de una secuencia particular de la salida.
- Dados los parámetros del modelo, encontrar la secuencia más probable de los estados ocultos que habrían podido generar una secuencia dada de la salida.
- Dado una secuencia de salida, ajustar los parámetros de modelo para calcular la probabilidad máxima de que esa secuencia se de o no.

Para el presente proyecto, dados los problemas planteados anteriormente y las funciones básicas de un HMM, el método consistirá de dos fases principales: *entrenamiento (problema 1) y detección de la señal de interés (problema 3)*.

Se asume que se tienen C =clases en el sistema que se supervisará y cada uno esta modelado por un HMM distinto y que se tienen conjuntos finitos de señales entrenamiento, donde cada patrón de entrenamiento constituye una secuencia O de la observación (o_1, o_2, \dots, o_T) de las características mas relevantes de los estados de sistema total. Con ello se explican los siguientes pasos:

- Paso 1: Construir un HMM λ para cada clase C que optimice la probabilidad de la secuencia del conjunto de observaciones con $P(O|\lambda)$ (la probabilidad de la secuencia de observaciones O dado el modelo λ).

El método a utilizar es el conocido como Baum-Welch, algoritmo de re-estimación iterativo para determinar el modelo a partir de una secuencia de observaciones de patrones de entrenamiento. También se utilizará el algoritmo *Forward-Backward*, algoritmo de sustitución para el cálculo de la probabilidad de una secuencia de observaciones dados los parámetros del modelo. Su desarrollo matemático puede estudiarse en [15].

- Paso 2: Proceder a detectar la señal de interés (en nuestro caso motosierras o disparos), donde el modelo con la mayor probabilidad es considerado como el mejor candidato para representar la condición dada, donde $P(O|\lambda_s)$ $1 \leq s \leq N$ y:

$$s^* = \arg \max [P(O|\lambda_s)] \quad (3.20)$$

Con los planteamientos anteriores, los tres problemas básicos de un HMM mencionados al inicio pondrían tener solución de la siguiente manera:

Solución del problema 1: [15]

Se quiere calcular la probabilidad de la secuencia de observaciones $O=O_1, O_2, \dots, O_T$ dado el modelo λ . La manera más directa de hacer esto es con la enumeración de cada posible secuencia del estado de longitud T (el número de observaciones).

$$Q = q_1 q_2 \cdots q_T \quad (3.21)$$

Donde q_1 es el estado inicial. La probabilidad de la secuencia de observaciones O para la secuencia de estados de 3.19 esta dada por:

$$P(O|Q, \lambda) = \prod_{t=1}^T P(O_t | q_t, \lambda = b_{q_1}(o_1) b_{q_2}(o_2) \dots b_{q_t}(o_t)) \quad (3.22)$$

donde se asume independencia estadística de observaciones. Además la probabilidad de la secuencia de estados es:

$$P(Q | \lambda) = \prod_{q_1} a_{q_1 q_2} a_{q_2 q_3} \dots a_{q_{t-1} q_t} \quad (3.23)$$

Inicialmente se plantea una solución en [15] poco eficiente, por lo que aquí se desarrolla el método de *Forward-Backward*, el cual considera una variable $\alpha_t(i)$ definida con el teorema de inducción matemática como:

$$\alpha_t(i) = P(O_1 O_2 \cdots O_t, q_t = S_i | \lambda) \quad (3.24)$$

la cual se entiende como la probabilidad de la secuencia parcial de observaciones $O=O_1, O_2, \dots, O_T$, hasta el tiempo t en el estado S dado el modelo λ . Este problema se puede resolver como sigue (ver explicación detallada en [15]):

- Inicialización:

$$\alpha_T(i) = \pi_i b_i(O_T), \quad 1 \leq i \leq N \quad (3.25)$$

- Inducción:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(O_{t+1}), \quad 1 \leq t \leq T-1 \quad (3.23)$$

$$1 \leq j \leq N$$

- Terminación:

$$P(O|\lambda) = \sum_{i=1}^N \alpha_T(i) \quad (3.24)$$

La figura 3-10 muestra la secuencia de operaciones requerida para el cálculo de la variable de *forward* α , donde el estado S_j puede se puede obtener desde los demás estados en un tiempo $t + 1$.

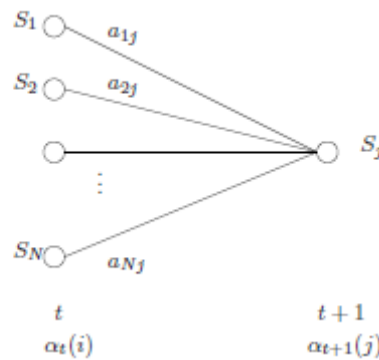


Figura 3- 12 Variable de forward α [15]

De manera similar se resuelve la variable de *backward*

- Inicialización:

$$\beta_T(i) = 1, \quad 1 \leq i \leq N \quad (3.25)$$

- Inducción:

$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(O_{t+1}) \beta_{t+1}(j), \quad (3.26)$$

$$t = T-1, T-2, \dots, 1, 1 \leq i \leq N$$

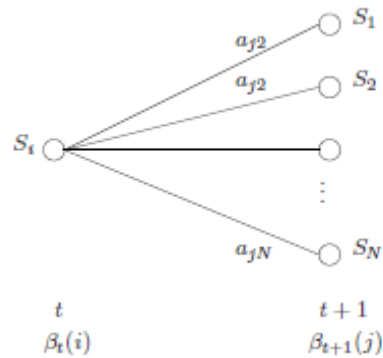


Figura 3- 13 Variable de backward β [15]

La Figura 3-13 muestra la secuencia de operaciones pero en esta ocasión para el cálculo de la variable *backward* β .

Solución al Problema 3:

Determinar un método para ajustar los parámetros del modelo del problema es el más difícil de resolver para los HMM, los cuales maximizan la probabilidad de la secuencia de observaciones dada por dicho modelo. No hay forma analítica de resolver el modelo que maximiza la probabilidad de la secuencia de observaciones,. De hecho, dando una secuencia de observaciones finitas como entrenamiento, no es

una forma óptima de estimar los parámetros del modelo. Sin embargo es aquí donde se utilizan los procedimientos iterativos como Baum-Welch [15, 23, 25,41], el cual al definir λ como el modelo actual y λ^* como el modelo reestimado, se puede llegar a estimar que:

- El modelo inicial define un punto critico en la función de probabilidad, en tal caso $\lambda=\lambda^*$.
- El modelo λ^* es más probable que el modelo λ y se encuentra un nuevo modelo con una probabilidad mayor que se produzca la secuencia de observaciones de entrenamiento hasta alcanzar un punto límite.

Capítulo 4. Procedimiento metodológico

4.1 Reconocimiento y definición del problema

En el proyecto se siguen etapas que permiten alcanzar los objetivos propuestos de acuerdo a los requerimientos establecidos por el asesor de la misma. La investigación es el punto inicial. A partir del conocimiento teórico y práctico se inicia con el desarrollo de los algoritmos y las pruebas que permitan evaluar la propuesta planteada de la extracción de información y clasificación de la señal, para obtener conclusiones relevantes. Las pruebas incluyen simulaciones, mediciones, depuraciones, entre otras.

Todas las especificaciones son importantes para ser tomadas en cuenta, por lo que es necesario desarrollar la documentación detallada para que se utilice de guía en el proyecto posterior de la solución de *hardware*.

4.2 Obtención y análisis de información

Para entender de forma adecuada el problema, se entrevista y se hacen las preguntas necesarias tanto a los profesores encargados, como a los estudiantes responsables de los diseños de las etapas anteriores, para recopilar datos y encontrar bases teóricas de las que se pueden derivar posibles soluciones.

El punto de partida importante es la investigación, la cual se realiza basada en bibliografía recomendada por los expertos, la búsqueda en Internet de *papers* dentro de las cuales se tiene acceso a bases de datos como la IEEE, entrevistas a instituciones y empresas que abarquen los temas estudiados y asistiendo a charlas o congresos que abarquen los temas de interés, ejemplo de esto es Iberchip.

Con la información obtenida se aplica el diseño de ingeniería para resolver el problema y con ello se realizarán las pruebas necesarias para comprobar que los

resultados cumplan con los objetivos propuestos y se darán las recomendaciones para la utilización del prototipo algorítmico a implementar.

4.3 Evaluación de las alternativas y síntesis de la solución

La investigación sobre todos los temas pertinentes para el proyecto, las consultas generadas a los encargados y el criterio personal dan la posibilidad de proponer la alternativa de solución al problema, analizando los algoritmos que se debían aplicar para métodos de reconocimiento de patrones.

Esto al lado de la experimentación y comprobación de resultados (simulaciones, mediciones, análisis continuo del diseño) llevaron a la mejor decisión de la solución que se desarrolla.

4.4 Implementación de la solución

Los algoritmos se implementan en MatLab, lenguaje que mediante el uso de bibliotecas específicas y *toolbox* existentes para *wavelets* y HMM, que permitieron la programación del prototipo. Además se utilizaron *Adobe Audition trial* (con licencia de prueba de 30 días) como *software* para la edición de las señales acústicas y Hisee (*software* libre) para la aplicación del método de sammon, útil para el cálculo de la variable K en K-means. Java 1.4.2 es la versión mínima necesaria para ejecutar dichos programas.

En el caso específico de los requerimientos de *hardware* para la ejecución del algoritmo diseñado son: Procesador Pentium III o superior, espacio suficiente en el disco (510 Mb) y 512 Mb de memoria RAM, recomendándose 1Gb, esto debido principalmente a los gráficos que se generan durante las pruebas.

4.5 Reevaluación y rediseño

Se plantearán las recomendaciones pertinentes, las cuales se basarán en el análisis del diseño propuesto y los resultados que se obtengan de éste. Así se busca lograr mejoras para la siguiente etapa del proyecto donde se implementará el *hardware* con

los algoritmos a diseñar para el reconocimiento de patrones de señales acústicas para motosierras y disparos.

Capítulo 5. Descripción detallada de la solución

El prototipo algorítmico diseñado utiliza procesamiento digital de señales basado en *wavelets* para la extracción de características de las señales acústicas y Modelos Ocultos de Markov para clasificarlas, descrita en capítulos anteriores.

Este diseño no se limita al reconocimiento de señales acústicas de disparos y motosierras, ya que dicho reconocimiento depende de las señales de entrada que sean entrenadas. Si el reconocimiento es dado o no, obedece básicamente a la cantidad de muestras ingresadas, la longitud de los datos de entrenamiento y de la secuencia de observaciones [39]. El rango de acción para la detección de los sonidos de interés varía de acuerdo con el entrenamiento dado a las señales como se explicará a continuación.

El desarrollo de la solución se divide en cuatro etapas:

- Extracción de características.
- Definición de las características básicas de los Modelos Ocultos de Markov.
- Entrenamiento de los modelos.
- Reconocimiento de las señales de interés.

5.1 Extracción de características

Debido a las propiedades de localización en tiempo y frecuencia, la transformada discreta de *wavelet* ha demostrado ser un punto de partida apropiado para la clasificación de las señales acústicas medidas.

Se presenta el esquema total del análisis de la señal en la figura 5-1, donde el conjunto de señales de entrenamiento S consisten en valores muestreados con condiciones como:

- Señales acústicas grabadas en vivo en el Parque Nacional Braulio Carrillo, teniendo un aproximado de 80 muestras (*Ver Apéndice A.2 para detalles*). Estas señales se capturaron a distancias de: 30m, 90m, 250m, 600m y tres diferentes ángulos: 0, 90 y 180 grados con respecto a la dirección del sonido grabado. La adquisición fue hecha en sonido monofónico, con un ancho de palabra de 16 bits @ 44.1kHz.
- Dichas señales fueron editadas para extraer de ellas únicamente la información que caracteriza ya sea a los disparos, motosierras o sonidos del ambiente, normalizando su amplitud al 100% mediante software guardando la fidelidad inicial del sonido.

Según Shaoxiang, Pittner, Kamarthi y Juditsky [29,32,34], no hay una receta general para la extracción de características con *wavelets*, ya que depende del tipo de señal, por lo que en este documento se propone inicialmente analizar cada señal acústica y se decide que el tamaño de las mismas sea de $D = 128K$ muestras para tener iguales condiciones de estudio. Ya editadas las señales, se experimenta con el ancho de los datos para el entrenamiento, dividiendo la señal específica en estudio en $G = 2$ (64K) y $G = 4$ (32K) partes. Además dicha señal tiene $F = \lceil \log_2 D \rceil$ bandas de frecuencia características.

A cada una de las divisiones hechas en la señal se le aplica la transformada discreta *wavelet* y luego de escoger un número determinado para F , se resuelve la ecuación (3.13) para calcular la energía promedio de los coeficientes *wavelet* de cada señal.

Los vectores resultantes contienen la energía promedio asociada a cada banda de frecuencia y estos son almacenados en una base de datos para compararlos mas adelante con las señales de evaluación.

Con estos vectores se forma una matriz H que contiene el número total de vectores discretos calculados, en la cual se emplea *K-means* para obtener el alfabeto de los modelos.

Para determinar la secuencia de observaciones \mathbf{O} , se compara la distancia euclidiana de cada vector de entrada dados en las matrices obtenidas de la energía promedio para cada señal, con los vectores del alfabeto uno a uno. De aquí se obtiene el índice del vector de entrada más cercano en la base de datos, al vector de energía en estudio. Para ir formando \mathbf{O} se escoge el índice para el que la distancia sea menor utilizando la ecuación (3.16). Esta cuantificación asigna un valor discreto entendible por el modelo oculto de Markov, como dicha observación.

Luego de este paso continúa el entrenamiento del modelo con Baum – Welch como se observa en la figura 5-1.

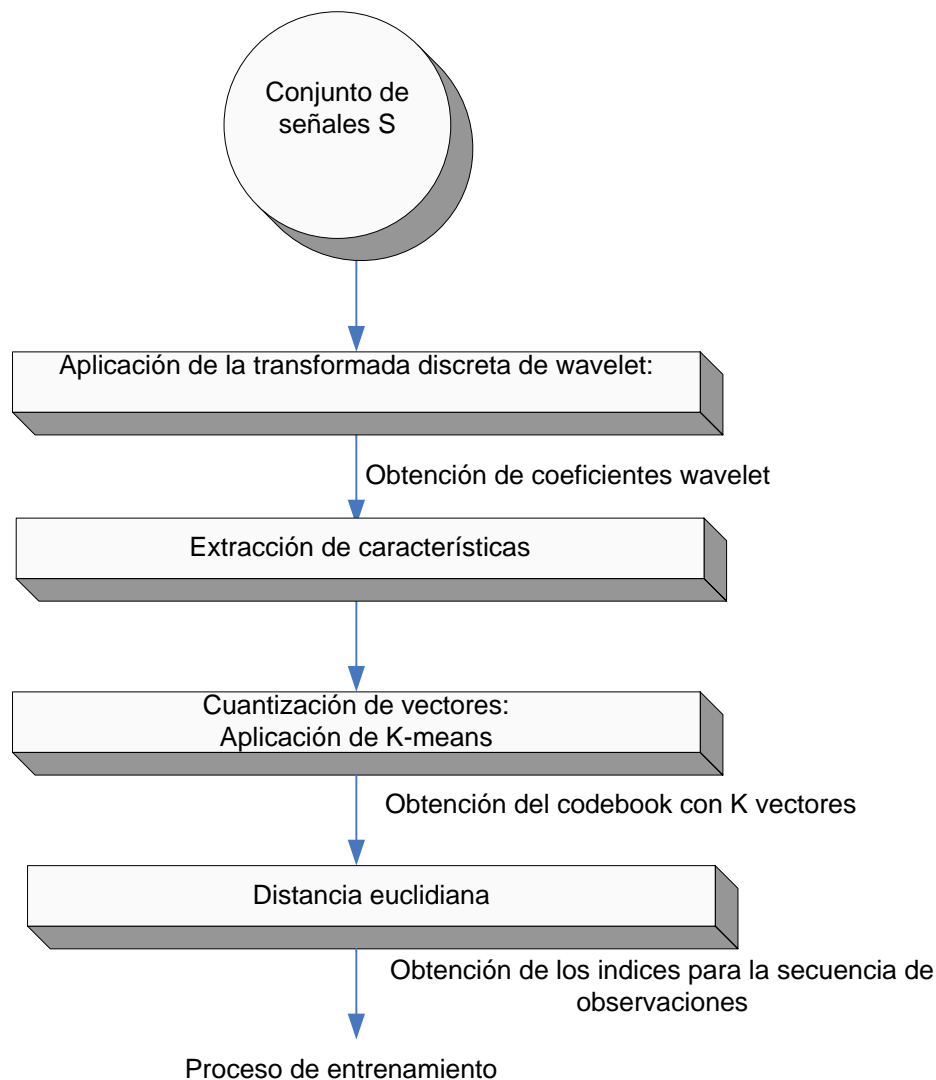


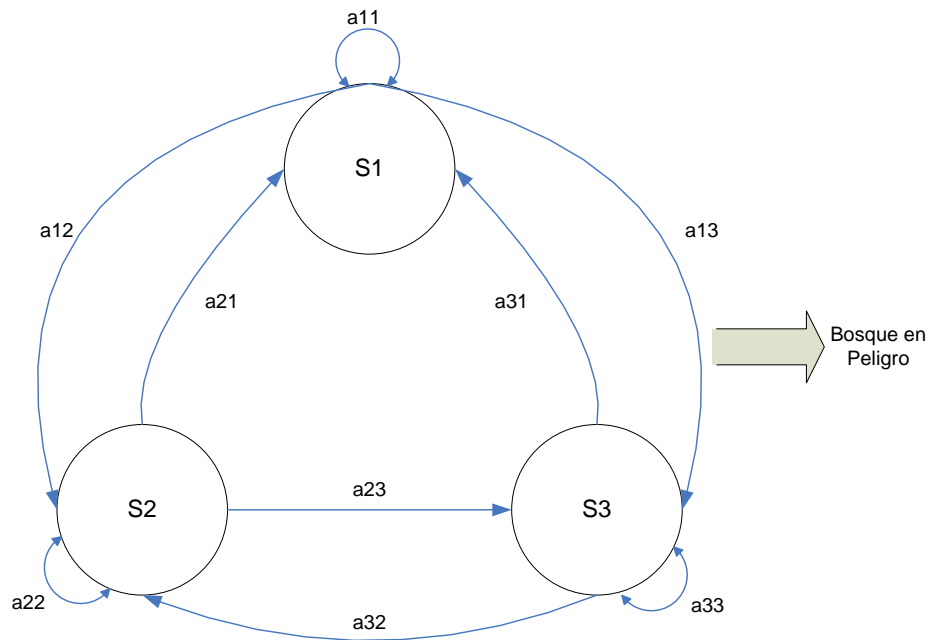
Figura 5- 1 Proceso de extracción de características de señales acústicas

5.2 Definición de las características básicas de los Modelos Ocultos de Markov

Se seleccionan Modelos Ocultos de Markov para desarrollar el prototipo algorítmico de reconocimiento de señales acústicas para la protección del ambiente, ya que pueden representar cualquiera de los símbolos observados [15].

Para el problema planteado se definen dos HMMs de primer orden, donde los patrones de entrada deberán ser asignados a las clases **peligro** si se detecta el sonido de una motosierra o disparo, o si se encuentra con sonidos naturales del

bosque entonces designa un ambiente **normal**. Para cada uno de ellos se consideran 3 estados N (S_1 , S_2 , S_3) como los observados en la figura 5-2, escogidos arbitrariamente ya que se desconocen cuántas posibilidades y cuáles podrían ser específicamente éstos en cualquier instante de tiempo. Los estados están interconectados de manera tal que cualquier estado oculto pueda alcanzar a otro.



(a)

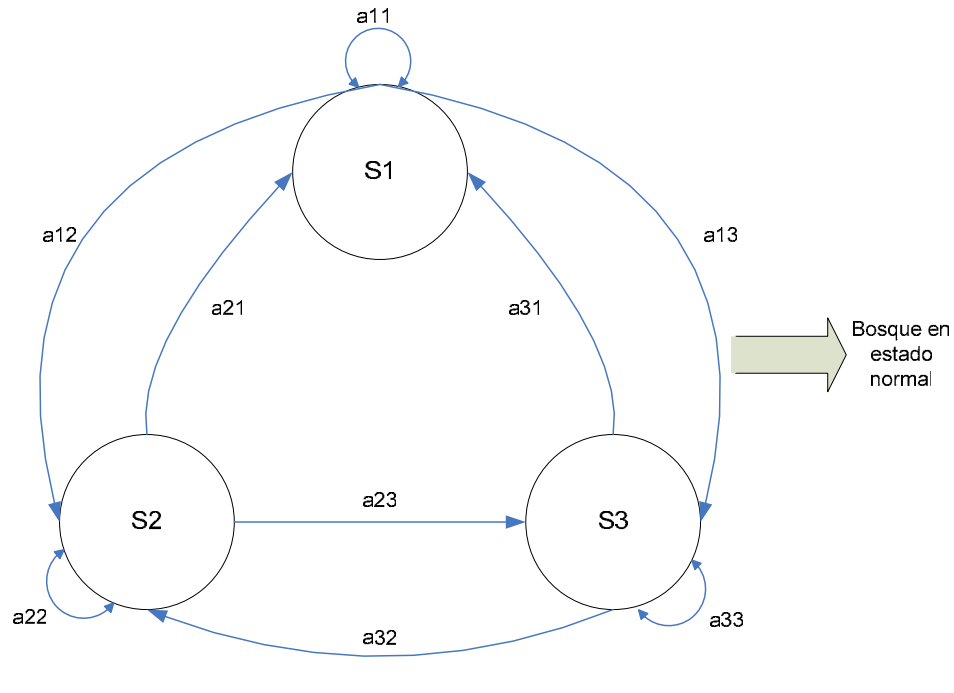


Figura 5- 2 HMM primer orden, 3 estados. (a) modelo de bosque en peligro (b) modelo de bosque en estado normal

El conjunto de observaciones \mathbf{V}_i , contiene el número de símbolos M_i donde M puede ser designado como peligro (disparos o motosierras) o normal:

$$\mathbf{V}_{\text{peligro}} = \{M_{1p}, M_{2p}, M_{3p}, \dots\}$$

$$\mathbf{V}_{\text{normal}} = \{M_{1n}, M_{2n}, M_{3n}, \dots\}$$

Cada clase es modelada por un HMM distinto, y para el entrenamiento de cada una de ellas se tiene un conjunto finito de señales, con las cuales se obtiene la secuencia de observaciones $\mathbf{O} = \{o_1, o_2, o_3, \dots, o_n\}$ con las representaciones apropiadas de las características de cada estado que se esperan reconocer.

5.3 Entrenamiento de los modelos

Al plantear el HMM para cada clase, el primer paso es estimar los parámetros iniciales del mismo, donde cada modelo esta definido por: $\lambda = [\mathbf{A}, \mathbf{B}, \pi]$.

Dichos parámetros iniciales son escogidos arbitrariamente y distribuidos de manera balanceada, donde \mathbf{A} es una matriz cuadrada de $S \times S$ (S representa el número de estados definidos en cada modelo):

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad (5.1)$$

\mathbf{B} que es la probabilidad de distribución de los símbolos de observación, esta determinada en sus columnas por el tamaño de C = alfabeto entrenado y sus filas por el número de estados S :

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & \dots & b_{1C} \\ b_{21} & b_{22} & b_{23} & \dots & b_{2C} \\ b_{31} & b_{32} & b_{33} & \dots & b_{3C} \end{bmatrix} \quad (5.2)$$

Finalmente π es la probabilidad de iniciar en un estado dado, por lo que se puede representar con un vector $\pi = [P_1, P_2, P_3]$.

Luego los parámetros planteados se re-estiman utilizando el método de *Baum-Welch*, esto para optimizar la secuencia de observaciones que se utiliza en el entrenamiento y encontrar la probabilidad máxima de la secuencia de observaciones \mathbf{O} dado el modelo ($P(\mathbf{O}|\lambda)$).

El algoritmo seguido para el entrenamiento se muestra en la figura 5.3:

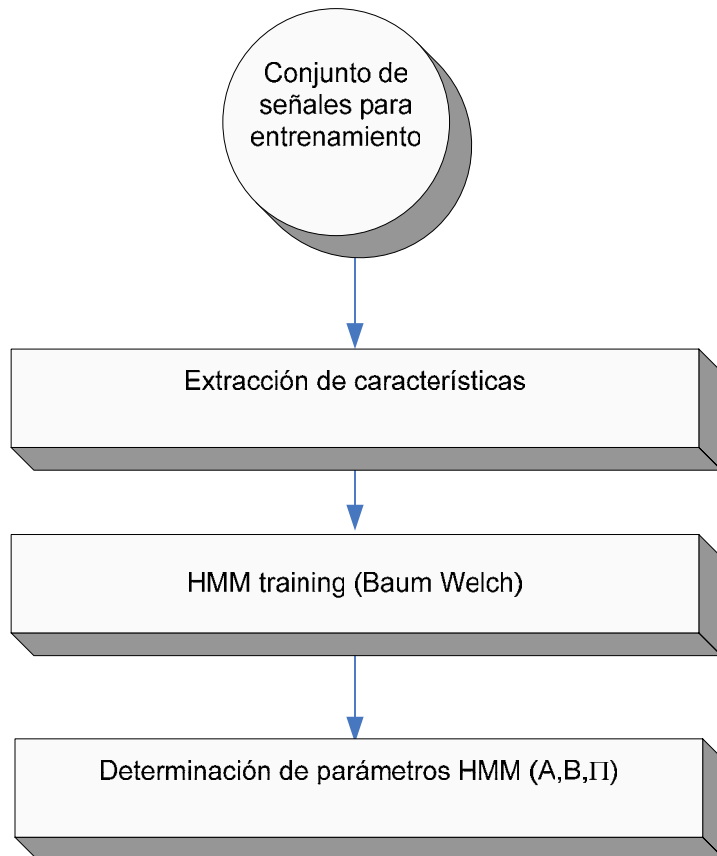


Figura 5- 3 Diagrama de flujo para el procedimiento de entrenamiento de HMM

5.4 Reconocimiento de señales

Después de haber calculado los coeficientes *wavelet* y desarrollado los vectores de cuantificación se calcula la máxima probabilidad de la secuencia de observaciones dada para que cada señal de entrada sea evaluada en cada uno de los modelos, como se ilustra en la figura 5.4:

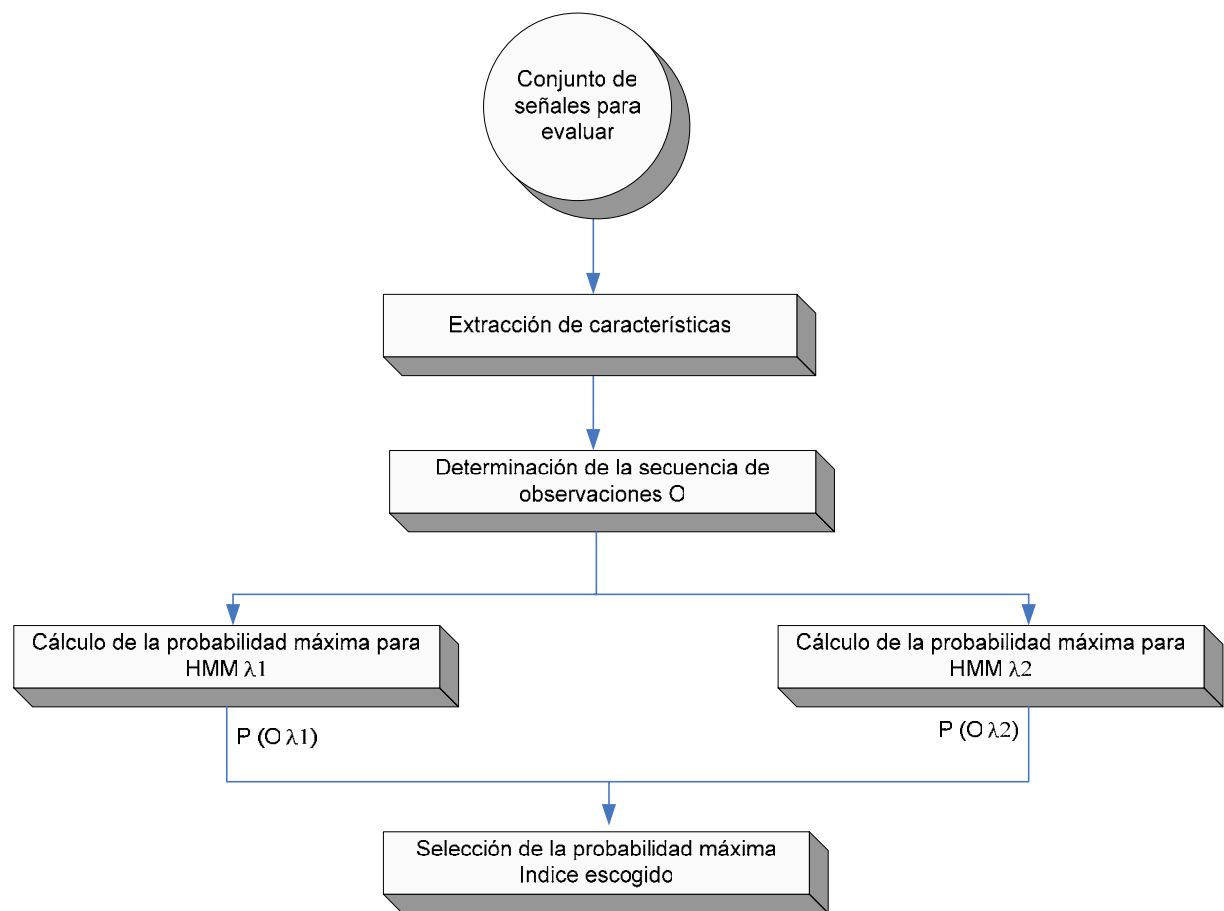


Figura 5- 4 Diagrama de bloques para el procedimiento de la condición de detección de la señal

El modelo con la máxima probabilidad es considerado el mejor candidato para representar la condición de la señal dada, matemáticamente dado por la ecuación (3.20). Computacionalmente se diseña con el algoritmo *Forward-Backward* visto en el capítulo 3.

Capítulo 6. Análisis de resultados

6.1 Extracción de características de la señal

Como parte de la experimentación se desarrollaron varios modelos para comparar su efectividad de reconocimiento, cambiando entre ellos la longitud de los datos de entrenamiento (cantidad de grupos de muestras analizadas por señal) y la longitud de la secuencia de observaciones (tamaño del alfabeto). Se seguirá un ejemplo para una misma señal en un modelo específico con ambas variantes mencionadas y con ello se ilustra el procedimiento y los resultados obtenidos.

El modelo a utilizar contiene un total de 80 señales entrenadas que incluyen sonidos de disparos, motosierras y sonido ambiente a las distancias y ángulos propuestos en el apéndice A.2.

Se tomará como ejemplo la señal de Carabina 22 – 30 metros – 0 grados, primero con un tamaño de 128 Kmuestras y longitud 2 (número de divisiones de 64Kmuestras cada una no traslapadas), luego la misma señal pero utilizando longitud 4 (número de divisiones de 32Kmuestras cada una no traslapadas). Ambos experimentos hechos mantienen las características de la señal necesarias para su posterior reconocimiento. Sus señales continuas se muestran en la figura 6-1

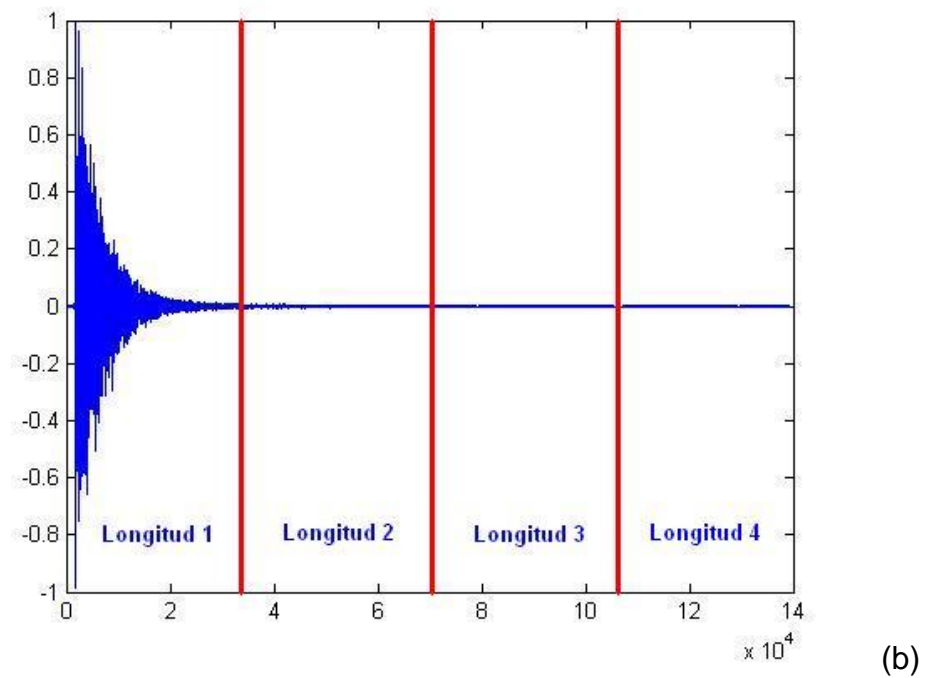
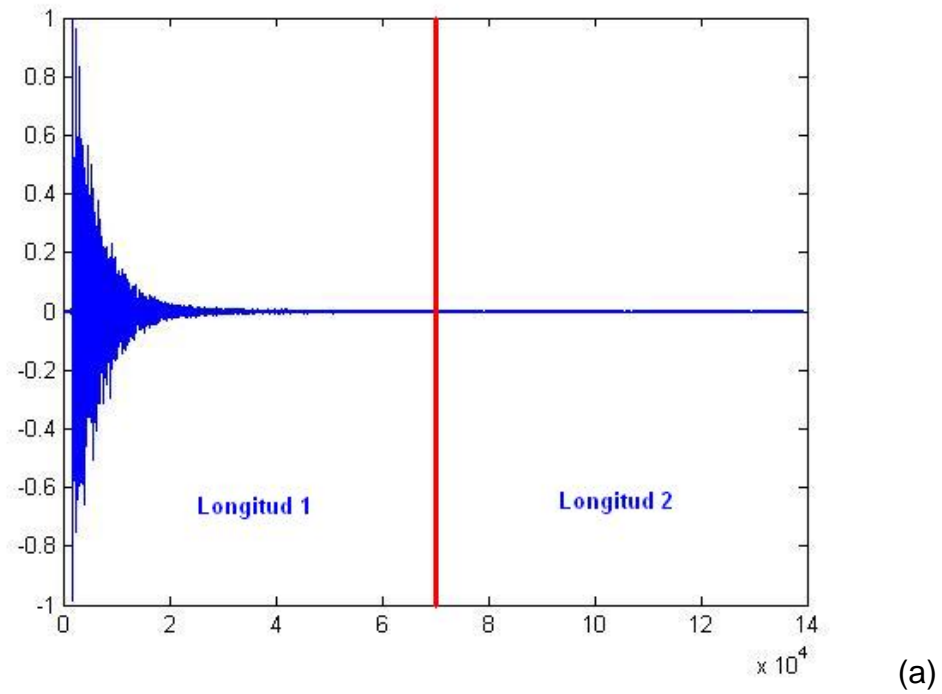
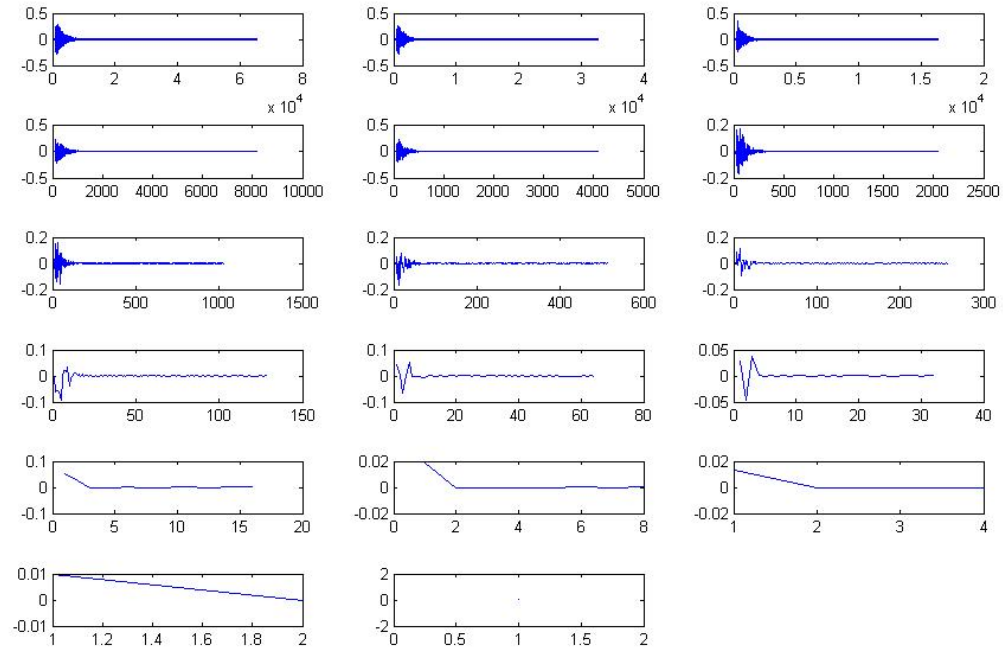
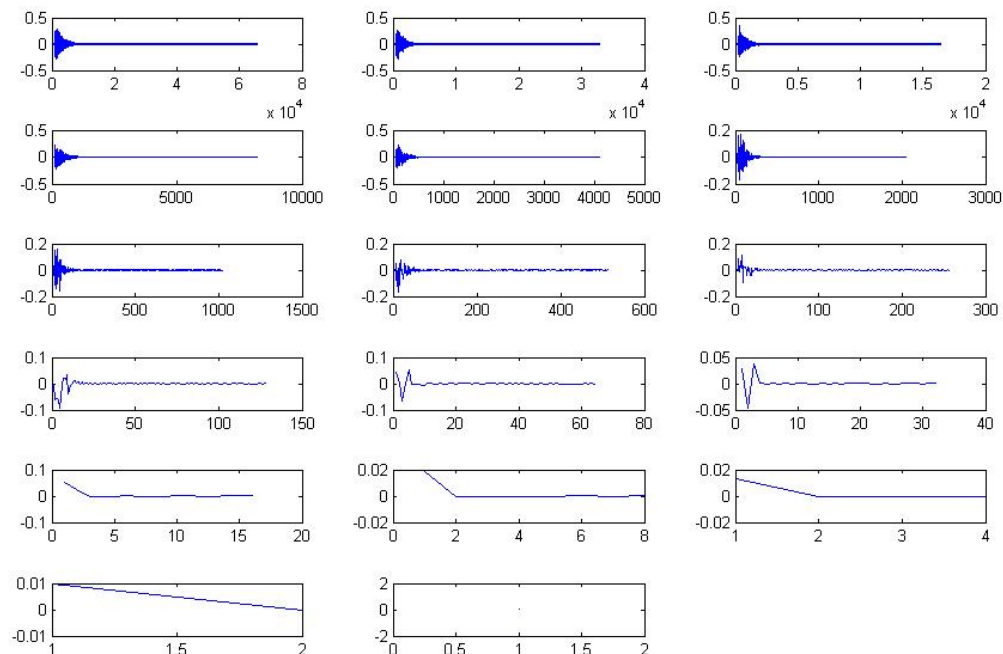


Figura 6- 1 Señal acústica (a) Señal de longitud 2. (b) Señal de longitud 4

Cuando se aplica la transformada *wavelet* se puede determinar el número de bandas de frecuencia para la señal con un máximo de $\log_2 N$, donde N es el número de muestras. Los coeficientes de energía para ambas se grafican en la figura 6-2, donde su eje x representa el número de muestras y el eje y su amplitud:



(a)

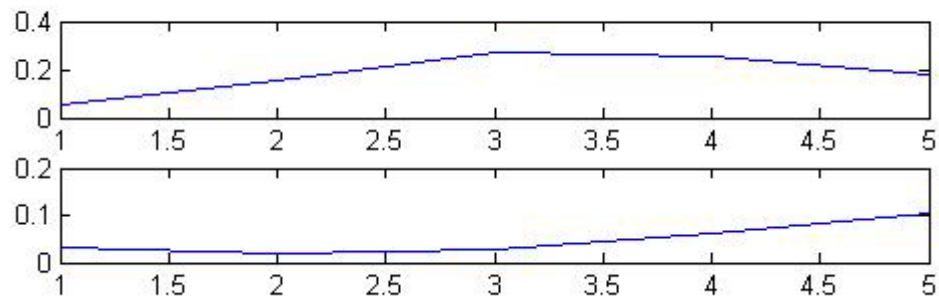


(b)

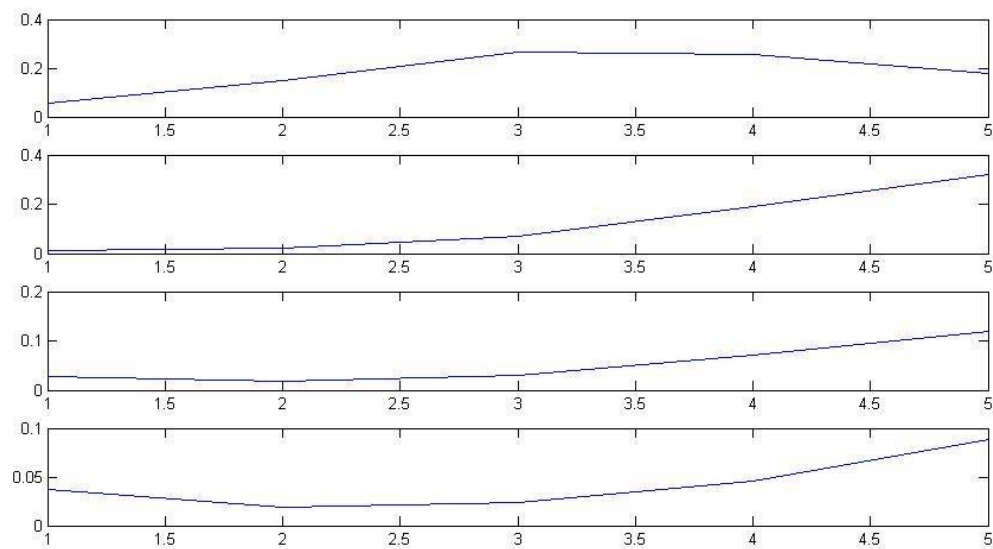
Figura 6- 2 Coeficientes de energía. (a) Señal de longitud 2. (b) Señal de longitud 4

Dadas sus características se escoge un número de 5 bandas de frecuencia. La salida ilustrada en la figura 6-2 de la transformada discreta de *wavelet* es una serie de vectores con coeficientes de aproximación que representan la señal variante en el tiempo y son descritos por la ecuación (3.8) la cual calcula el promedio de su energía.

Mediante la cuantificación del vector se representa de manera eficiente la información de tiempo – frecuencia de la señal, donde todas las características pueden ser representadas por un número pequeño y finito de vectores que forma más adelante el alfabeto. La representación gráfica de dichos vectores para el caso de la señal de ejemplo se muestra en la figura 6-3:



(a)



(b)

Figura 6- 3 Vectores de energía promedio para el arma Carabina 22, 30 metros, 0 grados. (a) Señal de longitud 2. (b) Señal de longitud 4

El eje x representa los niveles de energía escogidos, mientras que el eje y la amplitud de la señal. Nótese que se obtiene una cantidad de vectores de energía igual a la longitud escogida inicialmente para la señal. En los vectores se puede observar que las bandas sobresalientes son las de mayor frecuencia.

Siguiendo el proceso de extracción de la señal, cuando se crea la matriz H (matriz total), se le aplica *K-means* con el objetivo de obtener el alfabeto del modelo.

Para *K-means* es necesario un valor acertado de K que represente la distribución de los puntos espaciales adecuados para el modelo, con dimensión de columnas igual al número de bandas de frecuencia (en este caso 5), pero ilustrado en 2 dimensiones para comprenderlo. Para ello se utilizó el algoritmo de Sammon, donde se observa la distribución de los puntos en agrupaciones que se pueden unir y así determinar un K aproximado. En la figura 6-4 se muestra el resultado para el ejemplo en el caso de longitud 2 con K=48:

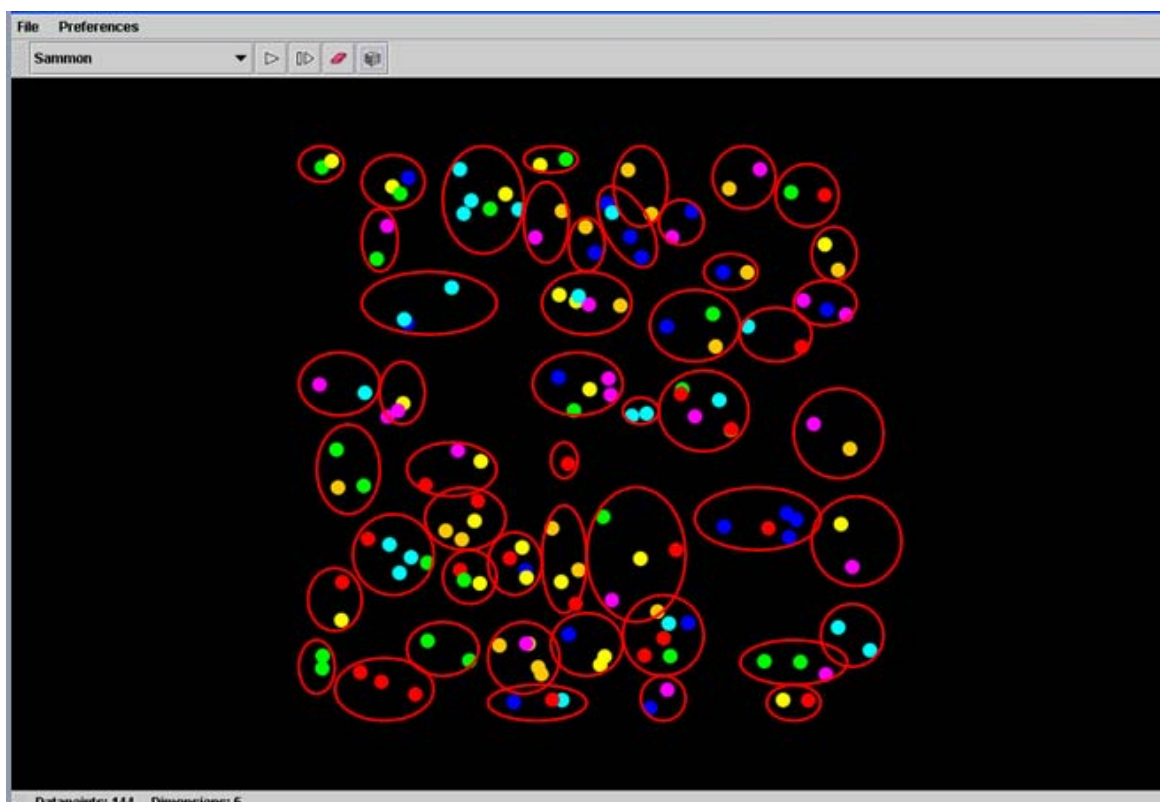


Figura 6- 4 Distribución espacial de 5 dimensiones, ilustrado en 2 dimensiones.

Con este valor, se aplica el algoritmo de *K-means* para generar el alfabeto de cada modelo. Luego la distancia euclidiana como se explicó en el capítulo 5 es calculada para crear dos secuencias de observaciones para el modelo de longitud 2 y dos secuencias de observaciones para el modelo de longitud 4 (una secuencia \mathbf{O}_1 para la clase de estado normal y una secuencia \mathbf{O}_2 para la clase de estado peligro), que posteriormente se entrenan. El número de índices determinados para cada secuencia es igual al número de vectores de entrada comparados en la distancia euclidiana, o lo que es igual el valor de K .

6.2 Definición de las características básicas de los Modelos Ocultos de Markov

Para detectar si el bosque esta en peligro o se encuentra en estado de ambiente normal, fueron plateados los modelos ocultos de Markov en el capítulo anterior con las siguientes características iniciales:

$$A = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} \quad (6.1)$$

$$\pi = [1,0,0] \quad (6.2)$$

Para cada clase inicial fue creada una matriz B : B_{Peligro} para el ambiente en peligro y B_{Normal} para un ambiente normal, donde dependiendo del modelo entrenado las columnas se definen por el tamaño del alfabeto.

$$B_{\text{peligro}, \text{normal}} = \begin{bmatrix} 1/C & 1/C & 1/C & 1/C & 1/C & 1/C & \dots \\ 1/C & 1/C & 1/C & 1/C & 1/C & 1/C & \dots \\ 1/C & 1/C & 1/C & 1/C & 1/C & 1/C & \dots \end{bmatrix} \quad \text{C: alfabeto, (6.3)}$$

6.3 Entrenamiento de los modelos

Al re- estimar dichos parámetros con el método de *Baum-Welch*, utilizando 100 iteraciones para el modelo de longitud 2 se obtienen dos matrices de 3x48 y para el modelo de longitud 4 dos matrices de 3x100, de la forma (6.3) y al observar sus resultados se comprueba que las características básicas de los HMM escogidos fueron acertadas, ya que las probabilidades de tener presente un sonido de peligro en la matriz de ambiente normal son iguales a cero y viceversa.

6.4 Evaluación

Para obtener iguales condiciones de evaluación se definen las siguientes condiciones:

- En cada modelo se evalúan las mismas 78 muestras asignadas.
- Todas las muestras entrenadas y evaluadas tienen un tamaño de 128 Kb.

Los tres modelos comparados en la tabla 6.1 tienen la siguiente cantidad de señales entrenadas:

- Longitud de observaciones a 30 m. 0 grados: 3 señales de ambiente y 6 señales de peligro.
- Longitud de observaciones a todas las distancias propuestas y 0 grados: 15 señales de ambiente y 20 señales de peligro.
- Longitud de observaciones a todas distancias propuestas y todos los grados mencionados: 15 señales de ambiente y 61 señales de peligro.

El proceso de entrenamiento de cada señal utilizando los requerimientos vistos en la sección 4.4 del informe tiene una duración aproximada de 1 minuto, mientras que su reconocimiento no sobrepasa 0.05 segundos.

Tabla 6. 1 Variación de la longitud de observaciones entrenadas con señales de longitud 2

	Longitud de observaciones 30 m, 0 grados		Longitud de observaciones todas distancias, 0 grados		Longitud de observaciones todas distancias, todos los grados	
	Modelo Ambiente Tranquilo	Modelo Ambiente Peligro	Modelo Ambiente Tranquilo	Modelo Ambiente Peligro	Modelo Ambiente Tranquilo	Modelo Ambiente Peligro
Modelo Ambiente Tranquilo	13	4	14	3	16	1
Modelo Ambiente Peligro	40	21	6	55	0	61
Probabilidad máxima de detección	52%	87%	90%	94%	100%	98%

Al variar la longitud de observaciones en el entrenamiento y evaluar un número constante de 78 muestras para todos los casos, se obtienen la probabilidades de detección de la tabla 6.1, donde para el primer caso por ejemplo se obtiene un 52% de detección para señales acústicas del ambiente tranquilo y un 87% para el ambiente en peligro. Se puede observar que conforme aumenta la longitud de las observaciones, así lo hace también la probabilidad de que el sonido sea reconocido por el modelo.

Tabla 6. 2 Variación de la longitud de observaciones evaluadas con señales de longitud 4.

	Longitud de observaciones 30 m, 0 grados		Longitud de observaciones todas distancias, 0 grados		Longitud de observaciones todas distancias, todos los grados	
	Modelo Ambiente Tranquilo	Modelo Ambiente Peligro	Modelo Ambiente Tranquilo	Modelo Ambiente Peligro	Modelo Ambiente Tranquilo	Modelo Ambiente Peligro
Modelo Ambiente Tranquilo	7	10	7	10	17	0
Modelo Ambiente Peligro	20	41	0	61	0	61
Probabilidad máxima de detección	67%	80%	100%	86%	100%	100%

Con los resultados de la tabla 6.2 donde además de la longitud de observaciones se aumentó la longitud de divisiones de la señal para analizarla (en este caso 4) mejora considerablemente la probabilidad de detección del sonido de interés, ya que se tienen mayor cantidad de datos procesados y entrenados en la secuencia de observaciones dada, lo que le da la oportunidad a los HMM de discriminar mejor el resultado final.

Los resultados de las tablas 6.1 y 6.2 se aprecian gráficamente en las figuras 6-5 y 6-6

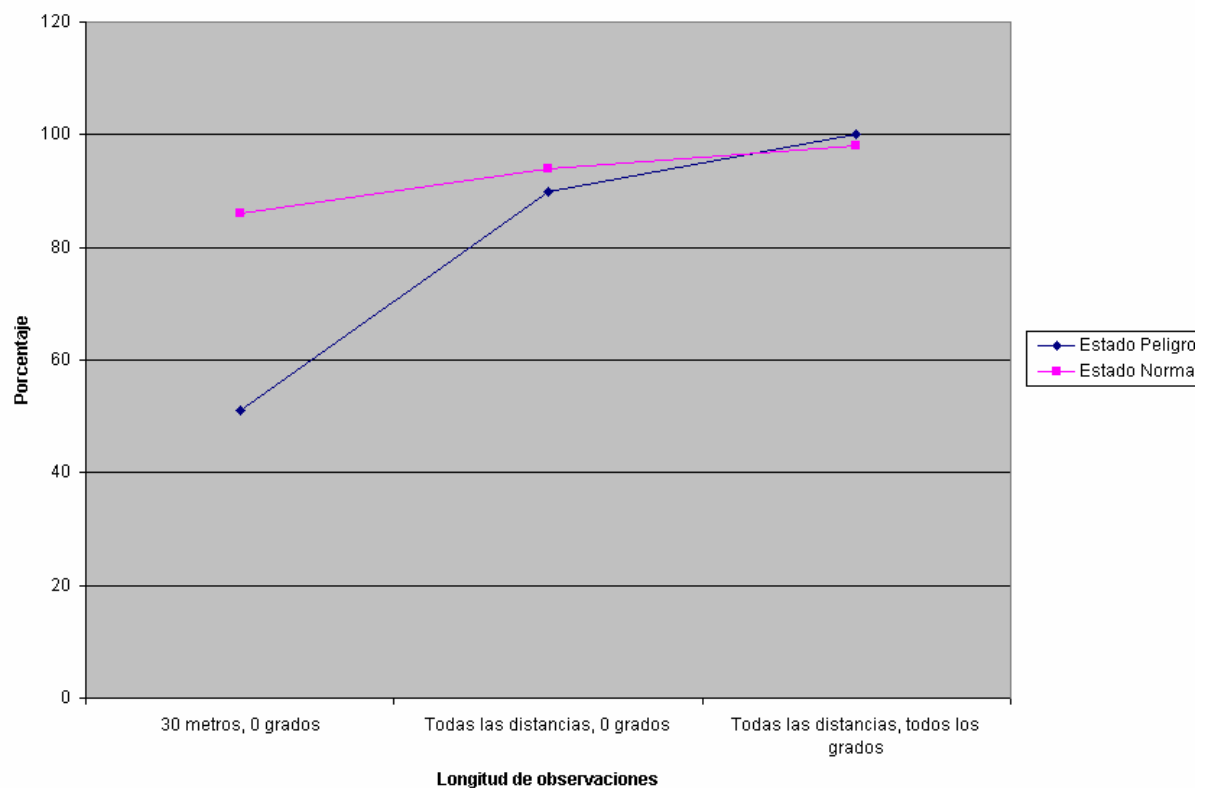


Figura 6- 5 Gráfico del cambio en porcentaje de acuerdo a la longitud de observaciones. Señales de entrenamiento longitud 2

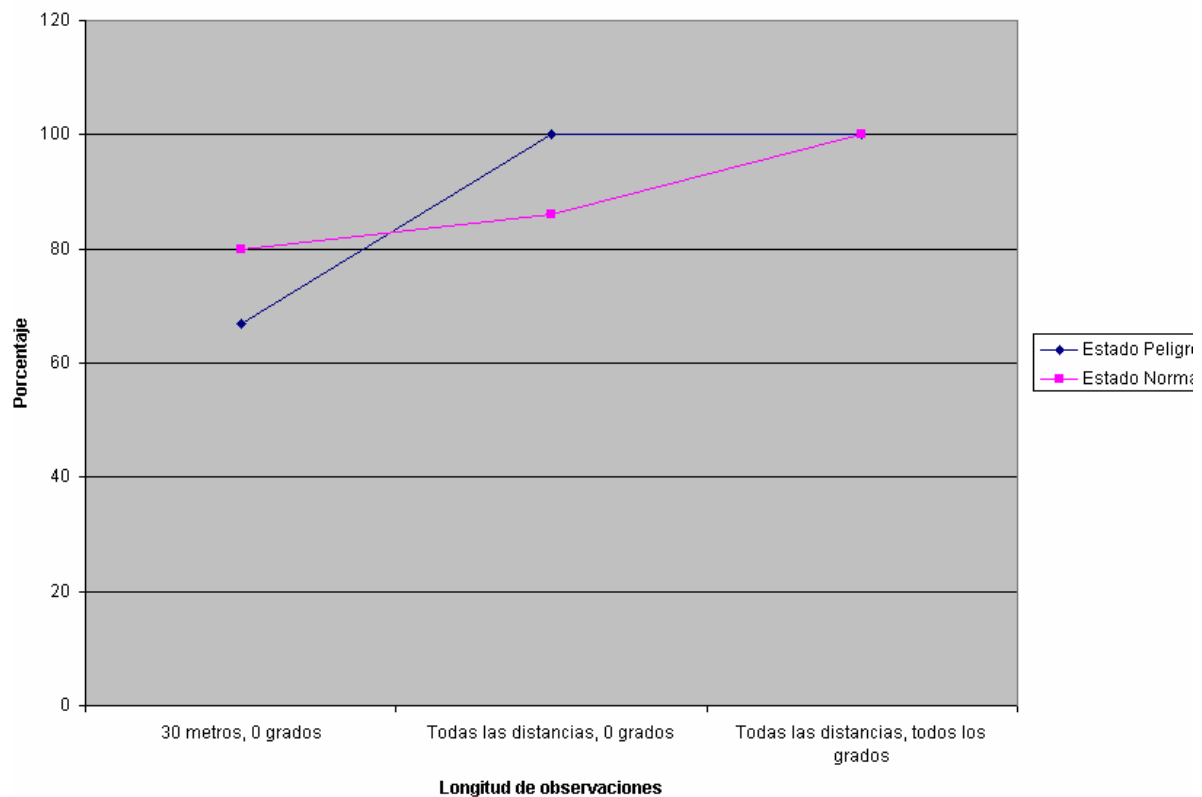


Figura 6- 6 Gráfico del cambio en porcentaje de acuerdo a la longitud de observaciones. Señales de entrenamiento longitud 4

Una de las metas para la implementación en *hardware* de este prototipo algorítmico es entrenar la mayor cantidad de sonidos para una detección exitosa de los mismos. Sin embargo se ha comprobado con las simulaciones, análisis e investigación hechos que inclusive con la cantidad de datos de señales de entrenamiento limitados los HMM generan una probabilidad alta de reconocer la señal de entrada, donde con solo 35 señales entrenadas (longitud 2, todas las distancias, 0 grados) para discriminar 5 tipos de armas, 2 motosierras, a 5 diferentes distancias y tres diferentes ángulos; además de los sonidos de la naturaleza que pueden interferir en el proceso, se logra una probabilidad de reconocimiento mayor al 90% y aumentando esta cantidad de señales de entrenamiento a 76 su reconocimiento es del 100%.

Capítulo 7. Conclusiones y recomendaciones

7.1 Conclusiones

- Se logró establecer una base datos amplia con los sonidos conocidos de motosierras y armas de fuego específicas y señales acústicas de una de las zonas boscosas del país para la verificación experimental de los algoritmos a diseñados. Dicha base de datos queda documentada para la etapa posterior del proyecto.
- La transformada wavelet es la herramienta matemática para el análisis de señales que cumplió con los requerimientos del proyecto, ya que permitió el diseño y aplicación de rápidos algoritmos para el cálculo computacional.
- En los modelos oculto de Markov, un aumento en longitud de la secuencia de la observación mejora el la probabilidad de reconocimiento de los modelos entrenados.
- La eficacia del sistema fue evaluada en diferentes condiciones (longitudes de entrenamiento y en la secuencia de observaciones), logrando una probabilidad mayor al 90%, aun con datos de entrenamiento limitados.
- El proyecto brinda una herramienta que permitiría detectar actividades ilegales en zonas protegidas de forma instantánea con tiempos de reconocimientos igual o menores a 0.05s, cumpliendo con los requerimientos especificados en 4.4, alertando de forma eficiente a las autoridades respectivas.

7.2 Recomendaciones

- Implementar los algoritmos de extracción y clasificación de señales en algún lenguaje de alto nivel que contenga librerías disponibles para su desarrollo y con plataformas que estén al alcance del usuario final en cuanto a licencias entre otros aspectos. Por ejemplo:

A. GHMM “*General Hidden Markov Model library (GHMM)*” [18]: Lenguaje C, gratis bajo la licencia LGPL [19], que implementa estructuras de datos y algoritmos para trabajar con HMM.

B. GSL “*GNU Scientific Library*” [18]. Métodos numéricos en lenguaje C/C++ [20]. Disponible de manera gratuita. Funciona para programar transformaciones wavelet y manipulaciones numéricas de datos intermedios entre otros.

C. SWIG [18]: Herramienta que permite conectar programas escritos en C y C++ con otros lenguajes de alto nivel [21] como Python [22], language de programación interpretado, con el objetivo de proveer una interfaz de alto nivel para el programador.

- Para posteriores etapas del proyecto se debería procurar la adquisición de la señal en sonido estereo para contar con información espacial implícita, y simplificar así la simulación de condiciones cuando se comprueben los sistemas utilizados.

Bibliografía

1. *Wavelet* [en línea]: <http://www.wikipedia.org/wiki/wavelet>. GNU free documentation license. [Consulta 9 Agosto 2005]
2. Escuela de Ingeniería Electrónica. Reseña histórica. [en línea]: <http://www.ietec.org/info.html>. [Consulta 8 Agosto 2005]
3. B, Jelinek. Generalized hierarchical search in the ISIP ASR system: Signals, Systems and Computers. Conference Record of the Thirty-Fifth Asilomar Conference on. Volumen 2: pp 1553-1556. Estados Unidos: IEEE, 2001.
4. Helversen , Dagmar von; Helversen, Otto von. Acoustic pattern recognition in a grasshopper: processing in the time or frequency domain?. Volumen 79: número 6. Berlin: *Biological cybernetics* , Febrero de 2004.
5. Leiva Rojas, Arnoldo. Diseño e implementación de la etapa de detección de disparo de armas de la red inalámbrica de telecomunicaciones para la protección de zonas protegidas. Proyecto de graduación. Cartago: Instituto Tecnológico de Costa Rica, 2005.
6. Salas Pérez, Walter. Diseño e implementación de un sensor para la detección de motosierras de la red inalámbrica de telecomunicaciones de la protección ambiental en el bosque. Proyecto de graduación. Cartago: Instituto Tecnológico de Costa Rica, 2005.
7. Advanced Wireless PowerCode™ Glass-Break Detector - MCT-501 [en línea]: <http://www.safe-mart.com/adwipogldemc.html> [Consulta 14 Agosto 2005]
8. [en línea]: ismir2002.ismir.net/proceedings/03-SP02-4.pdf [Consulta 14 Agosto 2005]
9. Berg, Tone. Pattern recognition [en línea]: <http://www.sintef.com> [Consulta 14 Agosto 2005]
10. Proakis, J. Manolakis, D. Tratamiento digital de señales. 3ra edición. Madrid: Prentice, Mayo 2000.
11. Robles, Edgar. Agenda para la competitividad de Costa Rica hacia el siglo XXI. [en línea]:

<http://www.incae.ac.cr/ES/clacds/investigacion/pdf/agendacostarica.pdf>
[Consulta 7 Agosto 2005]

12. *Assumptions in the theory of HMMs* [en línea]: <http://www.wikipedia.org/wiki/hmm>. [Consulta 9 Agosto 2005]
13. *Redes Neurales* [en línea]: <http://www.wikipedia.org/wiki/redes> [Consulta 9 Agosto 2005]
14. *Reinforcement learning* [en línea]: <http://leemon.com/papers/residual/residual.pdf> [Consulta 23 Agosto 2005]
15. Lawrence, R. *A tutorial on HMM and selected applications in speech recognition: Digital Object Identifier*. Volumen 77. No2, Estados Unidos: IEEE, febrero 2005.
16. Pirchio, F. *An acoustic surveillance unit for energy aware sensor network*. XII Workshop Iberchip: pp 191-194. San Jose, Costa Rica: 2006.
17. Sanchez, Alonso. *Predicción de fallas y caracterización de estado de uso para equipo de laboratorio utilizando clasificadores estadísticos*. Proyecto de graduación. Cartago: Instituto Tecnológico de Costa Rica, 2006.
18. *Free Software Foundation. Gnu lesser general public license* [en línea]: <http://www.gnu.org/copyleft/lesser.html>. [Consulta 1 Noviembre de 2006].
19. *Free Software Foundation. Gnu scientific library* [en línea]: <http://www.gnu.org/software/gsl/>. [Consulta 1 Noviembre de 2006].
20. *Simplified wrapper and interface generator. University of Chicago*. [en línea]: <http://www.swig.org/>. [Consulta 2 Noviembre de 2006].
21. Van Rossum, Guido. *Python programming language* [en línea]: <http://www.python.org/>. [Consulta 2 Noviembre de 2006].
22. Veeravalli, Anant. *A TUTORIAL ON USING HIDDEN MARKOV MODELS FOR PHONEME RECOGNITION: Digital Object Identifier*. Volumen, 20-22 pp: 154-157. Estados Unidos: IEEE, Marzo de 2005.
23. Mohamed, M; Gader, P. *Generalized hidden Markov models. Theoretical frameworks: Digital Object Identifier*. Volumen 8 pp: 67-81. Estados Unidos: IEEE, Febrero del 2000.

24. Lim, S; Clements, M. Pseudo-continuous hidden Markov modeling for automatic speech recognition: Digital Object Identifier. Volumen 2 pp: 482-488. Estados Unidos: IEEE, Abril de 1992.
25. Abdel-Galil, T. Disturbance classification using Hidden Markov Models and vector quantization. Volumen 20: pp:2129-2135. Julio 2005.
26. Sanches, I. Noise-compensated hidden Markov models. Volumen 8.: pp 533-540. IEEE, Setiembre del 2000.
27. Myers, C. Modeling chaotic systems with hidden Markov models. IEEE Volumen 4: pp 565-568. IEEE, Marzo del 1992.
28. Pikrakis, A. Classification of musical patterns using variable duration hidden Markov models. Volumen 14: pp 1795-1807. IEEE, Setiembre de 2006.
29. Hu Shaoxiang. Signal denoising using wavelet packet hidden Markov model. Volumen 2: pp 751- 755. IEEE, Junio de 2004.
30. Krishnamurthy, V. Algorithms for scheduling of hidden Markov model sensors. Volumen 5: pp 4818- 4819. IEEE, Diciembre de 2001.
31. Varga, A. Hidden Markov model decomposition of speech and noise. Volumen 2: pp 845- 848. IEEE, Abril de 1990.
32. Pittner, S. Feature extraction from wavelet coefficients for pattern recognition tasks. Volumen 21: pp 83-88. IEEE, Junio de 1999.
33. Couvreur, Christophe. Hidden Markov Models and their mixtures. Reporte presentado para obtener el titulo de matematica. Universite Catholique de Louvain: 1996.
34. Juditsky, A. Wavelets in identification. INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE: Setiembre de 1994.
35. Fernandez, Pablo. Procesamiento Digital de Señales Acústicas utilizando wavelets. Instituto de Matematica UACH.
36. Kamarthi, S; Pittner, S. Fast Fourier and Wavelet Transform for Flank Wear Estimation—A Comparison Mechanical Systems and Signal Processing, Volumen 11: pp 791-809. Noviembre de 1997.

37. Wang, L.; Mehrabi, M. G; Kannatey-Asibu, 'Tool Wear Monitoring in Machining Processes Through Wavelet Analysis. Volumen XXIX. NAMRI/SME: 2001.
38. Vector Quantization [en línea]: <http://www.geocities.com/mohamedqasem/vectorquantization/vq.html> [Consulta 15 Octubre 2005].
39. Wang, L; Mehrabi, M; Kannatey-Asibu, E. HMM based Tool wear monitoring in turning. *Department of Mechanical Engineering, Engineering Research Center for Reconfigurable Machining Systems, University of Michigan, Ann Arbor, MI 48109-2125*
40. Lee, Kai-Fu. Automatic Speech Recognition, Kluwer Academic Publishers.: 1989.
41. Duda, R;Hart,P. Pattern Classification. Segunda edición. India: editorial Wiley, 2000.
42. Dan E. Tamir; Chi-Yeon Park; Wmk-Sung Yoo'. Vector Quantization and Clustering: A Pyramid approach. Transactions on Communications, Volumen COM-28: pp 84-95. IEEE, Enero 1980.
43. Hernández, Néstor. Red inalámbrica de telecomunicaciones para la protección del ambiente. Cartago: Instituto Tecnológico de Costa Rica, 2004.
44. Brown, Judith. Feature dependence in the automatic identification of musical woodwind instruments. [en línea]: <http://scitation.aip.org/> [Consulta Febrero 2006].

Apéndices

A.1 Glosario, abreviaturas y simbología

off line: Proceso de análisis hecho fuera de línea, no en el preciso instante en que se toma la muestra.

HMM: *Hidden Markov Models*. Modelos ocultos de Markov.

DWT: Transformada discreta de *wavelet*.

FWT: Transformada rápida de *wavelet*

CWT: Transformada continua de *wavelet*.

FIR: Respuesta Infinita al Impulso:

IIR: Respuesta Finita al Impulso.

Clustering: Agrupar

Wavelets: Pequeñas ondas

A.2 Protocolos de medición

Base de datos de sonidos de bosque, disparos y motosierras

Captura de datos

Lunes 14 de agosto, 2006

Parque Nacional Braulio Carrillo – Tajo Santa Elena

Costa Rica

Profesores: Ing. Msc. Néstor Hernández Hostaller
Ing. Dr. Pablo Alvarado Moya
Ing. Forestal Marvin Castillo

Estudiante: Maria Gabriela Sáenz Meza

Guarda Parques: Rolando Vargas.

Sonidista: Nano Fernández.

Disparos: Esteban Rodriguez.

Materiales:

4 GPS

1 Grabador de sonidos en disco duro: Fostex PD-6 DVD location recorder

1 Micrófono Sennheiser súper direccional MKH 416

5 armas de fuego

1 motosierra

Los experimentos pretenden recolectar sonidos del bosque lluvioso/nuboso a tempranas horas de la mañana, acompañados de sonidos de disparos y motosierras, capturados con diferentes ángulos de incidencia, a diferentes distancias.

Se deben capturar los sonidos a diferentes distancias, aumentando estas hasta que no se pueda percibir más la señal, de ser posible. Deben fijarse límites máximos y mínimos de captura. Se proponen por tiempo y material disponible:

30m, 90m, 250m, 600m

Deben protocolarse las distancias reales utilizadas, preferiblemente con GPS, para contar con información que permita generar modelos sobre la distorsión de la señal

con la distancia. Con respecto a esto, deben protocolarse los niveles de ganancia utilizados en los micrófonos cuando se ajusten para reducir los errores de captura por saturación, de modo que se cuente con información suficiente para normalizar a posteriori todas las muestras a un mismo nivel.

Existen dos posibilidades básicas para la adquisición:

- Dejar el punto de captura fijo, desplazando las armas y motosierras
- Dejar el punto de los disparos fijos y desplazar el punto de adquisición.

Por la complejidad del equipo de adquisición y su protección contra la lluvia, se recomienda utilizar la versión 1. De ser posible, cada serie experimental debería repetirse para contar con al menos dos o más variantes de la misma serie de datos. Esto es importante para el entrenamiento de clasificadores.

La adquisición es hecha en sonido “mono” inicialmente, con un ancho de palabra de 16 bits @ 44.1kHz por canal deberá asegurar la calidad.

A una misma distancia, deben capturarse los datos con el micrófono a diferentes orientaciones entre 0° y 180° de la línea que une a los puntos de adquisición y fuentes del sonido. Se propone utilizar:

0°, 90°, 180°

Con las distancias y ángulos propuestos, se contarían con 12 muestras por cada dispositivo (arma de fuego o motosierra).

Además de los sonidos de los disparos y motosierras, son relevantes para el entrenamiento de clasificadores aquellos sonidos que podrían confundirse con ellos: relámpagos, animales, sonidos de ramas quebrándose, voces humanas en el bosque, etc., así como el sonido “natural” del bosque. De este modo, las muestras deben contener sonidos lo más variados posible.

Considerando el uso de 5 armas y 2 motosierras se tomarán un total de aproximadamente 72 muestras más todas aquellas de sonidos adicionales.

Hora: Legada 6 am

Coordenadas del punto de adquisición: Norte 10.05° - Oeste 84.02° - Altura 1638 m

Clima: Sin lluvia, nublado.

Observaciones: Grabación de 16 bits a 48 KHz.

Anotar en cada celda las coordenadas del punto de emisión.

	Distancia (m)	Coordenadas	0°	90°	180°
a	30	Norte: 10,053° Oeste: 84,02238° Altura: 1636 m	1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0
b	90	Norte: 10,05286° Oeste: 84,02316° Altura: 1629 m	1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0
c	250	Norte: 10,05084° Oeste: 84,02167° Altura: 1637 m	1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0
d	600	Norte: 10,05302° Oeste: 84,02769° Altura: 1746 m	1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0	1 2 3 4 5 6 7 8 9 0

Anotar identificación de cada dispositivo

	Moto- sierra	Arma	Identificación del dispositivo
1		x	Carabina 22
2		x	Revolver 32
3		x	Revolver 38

	Moto-sierra	Arma	Identificación del dispositivo
4		x	Pistola 9 mm
5		x	Escopeta 12
6	x		Motosierra

Anotar en cada tabla el índice temporal de cada muestra, es decir, una manera de localizar luego en todos los datos cada dispositivo con su ángulo y distancia.

Distancia/Coordenadas a (30m):

Ángulo Dispositivo	0°	90°	180°
1	Toma 2	Toma 3	Toma 4
2	Toma 5	Toma 6	Toma 7
3	Toma 10	Toma 11	Toma 12
4	Toma 13	Toma 14	Toma 15
5	Toma 16	Toma 17	Toma 18
6	Toma 20	Toma 21	Toma 22

Notas: Hora de inicio: 6:10 am. Toma 1: Prueba. Toma 8: es grabación de avión.
Toma 9: Nula. Toma 19: Pájaro - Ambiente

Distancia/Coordenadas a (90m)

Ángulo Dispositivo	0°	90°	180°
1	Toma 23	Toma 24	Toma 25
2	Toma 26	Toma 27	Toma 28
3	Toma 29	Toma 30	Toma 31
4	Toma 32	Toma 35	Toma 36
5	****	****	****
6	Toma 37-38	Toma 39	Toma 40

Notas: Hora inicio: 6:50am. Las tomas tienen bastante ambiente = -30dB. Toma 31: Pájaros. Toma 32: Se alarga ambiente y viento. Toma 34: Avioneta. Toma 39: Plus Ambiente. Archivos de motosierra identificados como M90. Escopeta 12 no se utiliza debido al faltante de municiones.

Distancia/Coordenadas a (250m)

Ángulo Dispositivo	0°	90°	180°
1	Toma 41	Toma 42	Toma 43
2	Toma 44	Toma 45	Toma 46
3	Toma 47	Toma 48	Toma 49
4	Toma 50	Toma 51	Toma 52
5	Toma 53	Toma 54	Toma 55
6	Toma 56	Toma 57	Toma 58

Notas: Bastante ruido de fondo. Hora inicio: 7:30 am. Disparos se realizan cerca de una pared. Las tomas de la escopeta 12 tienen un jet de fondo. Los sonidos mas críticos son los de la carabina 22 y motosierra a 180 grados. Toma 59: Ambiente. Archivos de disparos identificados como D250 y M250 para motosierra.

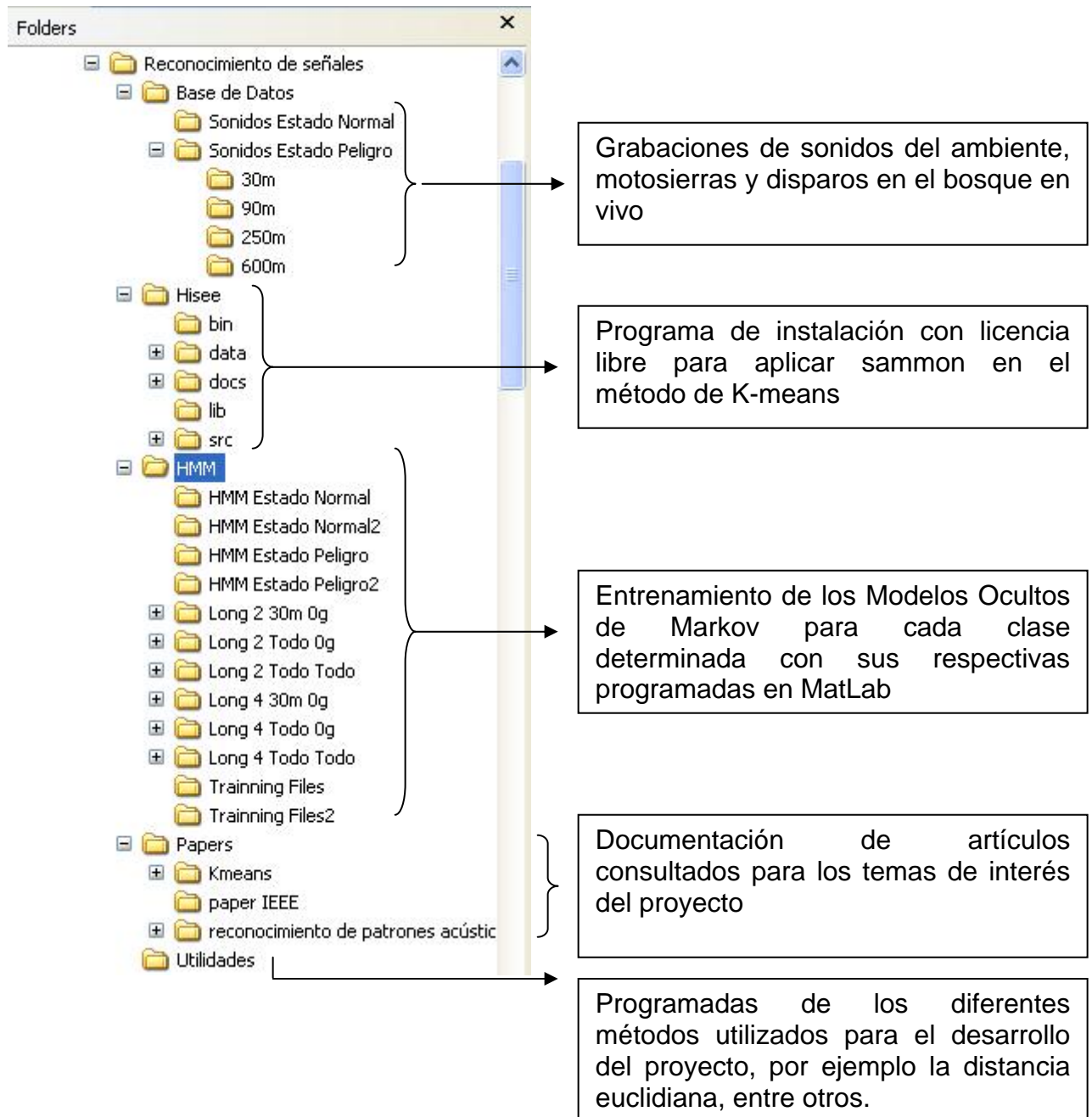
Distancia/Coordenadas a (600m)

Ángulo Dispositivo	0°	90°	180°
1	Toma 61	Toma 62	Toma 63
2	Toma 64	Toma 65	Toma 66
3	Toma 67	Toma 68	Toma 69
4	Toma 70	Toma 71	Toma 72
5	Toma 74	Toma 75	Toma ****
6	Toma 76	Toma 77	Toma 78

Notas: Hora inicio: 8:15 am. Archivos D600. Toma 60: invalida al igual que la 73 y 79.
Toma 80: lluvia.

Esta base de datos podrá ser encontrada en el CD de documentos finales de este proyecto.

A.3 Organización de documentación en CD



A.4 Información sobre la institución

El proyecto “Reconocimiento digital de patrones de sonidos de motosierras y disparos, el cual pertenece a la “Red inalámbrica de telecomunicaciones para la protección del ambiente”, se desarrolla en la Escuela de Ingeniería en Electrónica del Instituto Tecnológico de Costa Rica; creado el 10 de junio de 1971.[2]

La carrera de Electrónica se desarrolló mediante un Convenio de Cooperación Técnica entre los gobiernos de la República Federal de Alemania y de Costa Rica, con base en el cual, en octubre y noviembre de 1975 se suscribió un acuerdo para la creación del proyecto "Sección de Electrónica en el Instituto Tecnológico de Costa Rica", cuya misión sería la de “formar como técnicos en electrónica a bachilleres y obreros especializados con experiencia profesional que hayan aprobado el cuarto ciclo de enseñanza diversificada...” . Dicho proyecto se desarrolló a partir de enero de 1976 y por un lapso de tres años.

El Dipl. Ing. Gerd Lehm se constituyó en el primer director de la escuela y en conjunto con valiosos costarricenses puso en marcha el "Departamento de Electrónica". La república federal alemana, suministró los equipos técnicos de laboratorio, de enseñanza y materiales didácticos necesarios, y otorgó becas para la formación en dicho país de personal docente del Departamento de Electrónica; el Gobierno de Costa Rica, contrató el personal docente, administrativo y técnico, construyó aulas y laboratorios y dotó de mobiliario necesario.

En enero del 76 la carrera de Electrónica desarrolló el primer Plan de Estudios, que en término de tres años permitió la primer promoción, con el grado de Bachiller en Ingeniería Técnica en Electrónica. En 1979, mediante una ceremonia oficial, con participación de Autoridades de la Embajada de Alemania y del Instituto Tecnológico de Costa Rica, el Proyecto Departamento de Electrónica pasó a manos del Instituto.

Actualmente el programa está reconocido internacionalmente por el *CEAB* (*Canadian Engineering Accreditation Board*).

El marco del presente proyecto se realiza en el área procesamiento digital de señales, a cargo del Dr.-Ing. Pablo Alvarado Moya y tiene mucha relación con las áreas de procesamiento digital de señales, probabilidad y estadística. Es una investigación gestionada por el Msc.-Ing. Néstor Hernández y aprobada por la Vicerrectoría de Investigación y Extensión del Instituto Tecnológico de Costa Rica.